Mariano Suárez-Álvarez

Álgebra Lineal

21 de noviembre de 2016

Índice

Capítulo 1

Espacios vectoriales

§1. Cuerpos

1.1.1. Un ***cuerpo*** es un conjunto $\Bbbk$ dotado de dos operaciones

$$+ : \Bbbk \times \Bbbk \to \Bbbk, \qquad \cdot : \Bbbk \times \Bbbk \to \Bbbk,$$

a las que llamamos la ***suma*** y el ***producto*** de $\Bbbk$, respectivamente, que satisfacen las siguientes condiciones:

($\mathbf{K}_1$) la suma es asociativa: $(a+b)+c = a+(b+c)$ para cada $a, b, c \in \Bbbk$;

($\mathbf{K}_2$) la suma es conmutativa: $a+b = b+a$ para cada $a, b \in \Bbbk$;

($\mathbf{K}_3$) hay un elemento neutro para la suma: existe un elemento $0 \in \Bbbk$ tal que $a+0 = a$ y $0+a = a$ para todo $a \in \Bbbk$;

($\mathbf{K}_4$) todo elemento de $\Bbbk$ posee un opuesto aditivo: para todo $a \in \Bbbk$ existe un elemento $a' \in \Bbbk$ tal que $a+a' = 0$ y $a'+a = 0$;

($\mathbf{K}_5$) el producto es asociativo: $(a \cdot b) \cdot c = a \cdot (b \cdot c)$ para cada $a, b, c \in \Bbbk$;

($\mathbf{K}_6$) el producto es conmutativo: $a \cdot b = b \cdot a$ para cada $a, b \in \Bbbk$;

($\mathbf{K}_7$) hay un elemento neutro para el producto: existe un elemento $1 \in \Bbbk$ tal que $a \cdot 1 = a$ y $1 \cdot a = a$ para todo $a \in \Bbbk$;

($\mathbf{K}_8$) todo elemento de $\Bbbk$ distinto de 0 posee un inverso multiplicativo: para cada $a \in \Bbbk$ distinto de 0, existe $a' \in \Bbbk$ tal que $a \cdot a' = 1$ y $a' \cdot a = 1$;

($\mathbf{K}_9$) el producto se distribuye sobre la suma: $a \cdot (b+c) = a \cdot b + a \cdot c$ para cada $a, b, c \in \Bbbk$;

($\mathbf{K}_{10}$) los elementos neutros de la suma y del producto son distintos: $0 \neq 1$.

La condición ($\mathbf{K}_3$) afirma que existe un elemento neutro para la suma y, de hecho, implica que existe exactamente uno: en efecto, si 0 y $0'$ son dos elementos de $\Bbbk$ que son neutros para la suma, entonces

$$0 = 0 + 0' = 0'.$$

La primera de estas igualdades es consecuencia de que $0'$ es un elemento neutro para la suma, y la segunda de que 0 lo es. De manera similar, si a es un elemento de $\Bbbk$, la condición (K$_4$) nos dice que existe un elemento opuesto a a, esto es, un elemento $a' \in \Bbbk$ tal que $a + a' = 0$ y $a' + a = 0$, y este elemento está unívocamente determinado por a: si a'' es otro elemento opuesto a a, de manera que $a + a'' = 0$ y $a'' + a = 0$, entonces tenemos que

$$\begin{aligned}
a' &= a' + 0 && \text{por } (\mathrm{K}_3)\\
&= a' + (a + a'') && \text{porque } a'' \text{ es un opuesto a } a\\
&= (a' + a) + a'' && \text{por } (\mathrm{K}_1)\\
&= 0 + a'' && \text{porque } a' \text{ es un opuesto a } a\\
&= a''. && \text{por } (\mathrm{K}_3)\text{, otra vez.}
\end{aligned}$$

Como no hay entonces ambigüedad posible, escribiremos desde ahora en adelante $-a$ al elemento opuesto de a.

Procediendo de la misma forma, es fácil ver que el elemento 1 que es neutro para el producto cuya existencia afirma la condición (K$_7$) es uno sólo y que el inverso de todo elemento distinto de 0 que nos da la condición (K$_8$) está bien determinado. Gracias a esto último, podemos escribir sin ninguna ambigüedad a^{-1} al inverso de un elemento $a \in \Bbbk \setminus \{0\}$.

1.1.2. Ejemplos.

(a) Los conjuntos $\mathbb{Q}$, $\mathbb{R}$ y $\mathbb{C}$ de los números racionales, reales y complejos, respectivamente, dotados de sus operaciones usuales de suma y producto, son cuerpos.

(b) Sea $\mathbb{Q}(\sqrt{2})$ el conjunto de todos los números reales de la forma $a + b\sqrt{2}$ con $a, b \in \mathbb{Q}$. Si x e y son dos elementos de $\mathbb{Q}(\sqrt{2})$, entonces tanto la suma $x + y$ como el producto xy también están en $\mathbb{Q}(\sqrt{2})$: en efecto, si $x = a + b\sqrt{2}$ e $y = c + d\sqrt{2}$, con $a, b, c, d \in \mathbb{Q}$, entonces

$$x + y = (a + c) + (b + d)\sqrt{2}, \qquad xy = (ac + 2bd) + (ad + bc)\sqrt{2}$$

y es claro que los números $a + c$, $b + d$, $ac + 2bd$ y $ad + bc$ pertenecen a $\mathbb{Q}$. Esto nos dice que podemos restringir las operaciones de $\mathbb{R}$ a $\mathbb{Q}(\sqrt{2})$, obteniendo de esta forma operaciones

$$+ : \mathbb{Q}(\sqrt{2}) \times \mathbb{Q}(\sqrt{2}) \to \mathbb{Q}(\sqrt{2}), \qquad \cdot : \mathbb{Q}(\sqrt{2}) \times \mathbb{Q}(\sqrt{2}) \to \mathbb{Q}(\sqrt{2}).$$

Estas operaciones hacen de $\mathbb{Q}(\sqrt{2})$ un cuerpo. La verificación de todas las condiciones de la definición 1.1.1 es inmediata salvo la de (K$_8$), que es consecuencia de la siguiente observación: si x es un elemento no nulo de $\mathbb{Q}(\sqrt{2})$, de manera que existen $a, b \in \mathbb{Q}$ no simultáneamente nulos tales que $x = a + b\sqrt{2}$, entonces

$$x^{-1} = \frac{a}{a^2 + 2b^2} - \frac{b}{a^2 + 2b^2}\sqrt{2}.$$

(c) El conjunto $\mathbb{F}_2 = \{0,1\}$ es un cuerpo si lo dotamos de las operaciones de suma y producto dadas por

$+$	0	1
0	0	1
1	1	0

$\cdot$	0	1
0	0	0
1	0	1

(d) El conjunto $\mathbb{F}_3 = \{0,1,2\}$ es un cuerpo si lo dotamos de las operaciones de suma y producto dadas por

$+$	0	1	2
0	0	1	2
1	1	2	0
2	2	0	1

$\cdot$	0	1	2
0	0	0	0
1	0	1	2
2	0	2	1

(e) Más generalmente, se puede mostrar que si p es un número primo y $n \in \mathbb{N}$, entonces existe un cuerpo $\mathbb{F}_{p^n}$ con p^n elementos, que éste es esencialmente único — llamamos a estos cuerpos ***cuerpos de Galois***— y que los cuerpos que se obtienen de esta forma son los únicos cuerpos finitos.

(f) Si X es una variable y $\Bbbk$ es un cuerpo, el conjunto $\Bbbk[X]$ de los polinomios con coeficientes en $\Bbbk$ en la variable X no es un cuerpo cuando lo dotamos de sus operaciones usuales de suma y producto: por ejemplo, el polinomio X es no nulo y sin embargo no posee en $\Bbbk[X]$ inverso multiplicativo. En cambio, en conjunto $\Bbbk(X)$ de las ***funciones racionales***, esto es, el de de las fracciones $p(X)/Q(X)$ con p y q elementos de $\Bbbk[X]$ y q no nulo, es un cuerpo con respecto a sus operaciones usuales.

(g) Los conjuntos $\mathbb{N}$, $\mathbb{N}_0$ y $\mathbb{Z}$ de los enteros positivos, de los enteros no negativos y de los enteros, respectivamente, no son cuerpos cuando los dotamos de sus operaciones usuales: en $\mathbb{N}$ no hay un elemento neutro para la suma, en $\mathbb{N}_0$ no todo elemento posee un opuesto, en $\mathbb{Z}$ no todo elemento no nulo posee un inverso. ◇

§2. Espacios vectoriales

1.2.1. Fijemos un cuerpo $\Bbbk$. Un ***espacio vectorial sobre*** $\Bbbk$ es un conjunto V dotado de dos operaciones

$$+ : V \times V \to V, \qquad \cdot : \Bbbk \times V \to V$$

que llamamos la ***suma*** de V y la ***multiplicación escalar***, que satisfacen las siguientes condiciones:

($\mathbf{V}_1$) la suma es asociativa: $(x+y)+z = x+(y+z)$ para cada $x, y, z \in V$;

($\mathbf{V}_2$) la suma es conmutativa: $x+y = y+x$ para cada $x, y \in V$;

(**V**$_3$) hay un elemento neutro para la suma: existe un elemento $0_V \in V$ tal que $x + 0_V = x$ y $0_V + x = x$ para todo $x \in V$;

(**V**$_4$) todo elemento de V posee un opuesto: para todo $x \in V$ existe un elemento $x' \in V$ tal que $x + x' = 0_V$ y $x' + x = 0_V$;

(**V**$_5$) la multiplicación escalar es asociativa: $a \cdot (b \cdot x) = (a \cdot b) \cdot x$ para todo $a, b \in \Bbbk$ y todo $x \in V$;

(**V**$_6$) la multiplicación escalar es unitaria: $1 \cdot x = x$ para todo $x \in V$;

(**V**$_7$) la multiplicación escalar se distribuye sobre la suma de $\Bbbk$ y sobre la suma de V:

$$a \cdot (x + y) = a \cdot x + a \cdot y \text{ para todo } a \in \Bbbk \text{ y todo } x, y \in V;$$
$$(a + b) \cdot x = a \cdot x + b \cdot x \text{ para todo } a, b \in \Bbbk \text{ y todo } x \in V.$$

Estas condiciones implican que hay en V un único elemento neutro para la suma: si 0_V y $0'_V$ son dos elementos neutros, se tiene que $0_V = 0_V + 0'_V = 0'_V$. Decimos que es el ***cero*** de V. De forma similar, si x es un elemento de V, entonces hay un único elemento opuesto a x: si x' y x'' son dos elementos opuestos a x, entonces

$$x' = x' + 0_V = x' + (x + x'') = (x' + x) + x'' = 0_V + x'' = x'',$$

y esto nos permite escribir sin ambigüedad $-x$ para denotar al elemento opuesto de x.

1.2.2. Si $\Bbbk$ es un cuerpo y V es un espacio vectorial sobre $\Bbbk$, llamamos a los elementos de $\Bbbk$ ***escalares*** y a los elementos de V ***vectores***. Si $\Bbbk = \mathbb{R}$, decimos generalmente que V es un ***espacio vectorial real*** en lugar de que es un ***espacio vectorial sobre*** $\mathbb{R}$ y si $\Bbbk = \mathbb{C}$, que es un ***espacio vectorial complejo***.

1.2.3. Ejemplos.

(a) Sea $n \in \mathbb{N}$, sea $\Bbbk$ un cuerpo y sea $\Bbbk^n$ el conjunto de las n-uplas de elementos de $\Bbbk$, que consideraremos *siempre* como columnas. Si dotamos a $\Bbbk^n$ de las operaciones $+ : \Bbbk^n \times \Bbbk^n \to \Bbbk^n$ y $\cdot : \Bbbk \times \Bbbk^n \to \Bbbk^n$ definidas de manera que

$$(x_1, \dots, x_n)^{\mathrm{t}} + (y_1, \dots, y_n)^{\mathrm{t}} = (x_1 + y_1, \dots, x_n + y_n)^{\mathrm{t}},$$
$$a \cdot (x_1, \dots, x_n)^{\mathrm{t}} = (a \cdot x_1, \dots, a \cdot x_n)^{\mathrm{t}}$$

para cada par de elementos $(x_1, \dots, x_n)^{\mathrm{t}}$ e $(y_1, \dots, y_n)^{\mathrm{t}}$ de $\Bbbk^n$ y cada $a \in \Bbbk$, entonces $\Bbbk^n$ es un espacio vectorial sobre $\Bbbk$.

(b) Más generalmente, sean $m, n \in \mathbb{N}$, sea $\Bbbk$ un cuerpo y sea $\mathsf{M}_{n,m}(\Bbbk)$ el conjunto de las matrices de n filas y m columnas con entradas en $\Bbbk$. El conjunto $\mathsf{M}_{n,m}(\Bbbk)$ es un espacio vectorial sobre $\Bbbk$ con las operaciones de suma $+ : \mathsf{M}_{n,m}(\Bbbk) \times \mathsf{M}_{n,m}(\Bbbk) \to \mathsf{M}_{n,m}(\Bbbk)$ y de multiplicación por escalares $\cdot : \Bbbk \times \mathsf{M}_{n,m}(\Bbbk) \to \mathsf{M}_{n,m}(\Bbbk)$ usuales:

- si $a = (a_{i,j})$ y $b = (b_{i,j})$ son dos elementos de $\mathsf{M}_{n,m}(\Bbbk)$, entonces la suma $a + b$ es la matriz $(c_{i,j})$ con $c_{i,j} = a_{i,j} + b_{i,j}$ para cada $i \in [\![n]\!]$ y $j \in [\![m]\!]$; y
- si $\lambda \in \Bbbk$ es un escalar y $a = (a_{i,j}) \in \mathsf{M}_{n,m}(\Bbbk)$, entonces $\lambda \cdot a$ es la matriz $c = (c_{i,j})$ con $c_{i,j} = \lambda a_{i,j}$ para cada $i \in [\![n]\!]$ y $j \in [\![m]\!]$.

Notemos que si $m = 1$, el espacio $\mathsf{M}_{n,1}(\Bbbk)$ coincide con el espacio $\Bbbk^n$ que describimos en el ejemplo anterior. ◇

1.2.4. Ejemplos.

(a) Sea X un conjunto no vacío y sea $\Bbbk$ un cuerpo. El conjunto $\Bbbk^X$ de todas las funciones $X \to \Bbbk$ es un espacio vectorial sobre $\Bbbk$ con respecto a las operaciones

$$+ : \Bbbk^X \times \Bbbk^X \to \Bbbk^X, \qquad \cdot : \Bbbk \times \Bbbk^X \to \Bbbk^X$$

tales que

$$(f+g)(x) = f(x) + g(x),$$
$$(a \cdot f)(x) = a \cdot f(x)$$

para cada $f, g \in \Bbbk^X$, cada $a \in \Bbbk$ y cada $x \in X$.

(b) Más generalmente, si X es un conjunto no vacío y V es un espacio vectorial sobre un cuerpo $\Bbbk$, el conjunto $W = V^X$ de todas las funciones $X \to V$ es un espacio vectorial sobre $\Bbbk$ con respecto a las operaciones

$$+ : V^X \times V^X \to V^X, \qquad \cdot : \Bbbk \times V^X \to V^X$$

tales que

$$(f+g)(x) = f(x) + g(x),$$
$$(a \cdot f)(x) = a \cdot f(x)$$

para cada $f, g \in V^X$, cada $a \in \Bbbk$ y cada $x \in X$. ◇

1.2.5. Ejemplos.

(a) Sea $\Bbbk$ un cuerpo y sea $V = \Bbbk[X]$ el conjunto de los polinomios en la variable X con coeficientes en $\Bbbk$. Si dotamos a V de la operaciones usuales de suma de polinomios y de multiplicación por escalares, entonces V es un espacio vectorial.

(b) Sea $V = \{*\}$ un conjunto con un único elemento, que denotamos $*$. Hay exactamente una función $+ : V \times V \to V$ y hay exactamente una función $\cdot : \Bbbk \times V \to V$, y es fácil verificar que V es un espacio vectorial con respecto a esas operaciones. El cero de V es $*$ y entonces podemos escribir $V = \{0_V\}$. Llamamos a este espacio vectorial un ***espacio vectorial nulo*** y lo escribimos, cuando esto no dé lugar a confusiones, simplemente 0.

(c) Si K es un cuerpo y $\Bbbk \subseteq K$ es un subcuerpo de K, entonces K es un espacio vectorial sobre $\Bbbk$ con respecto a la operación de suma $+ : K \times K \to K$ de K y a la multiplicación $\cdot : \Bbbk \times K \to K$ por elementos de $\Bbbk$ que se obtiene restringiendo la multiplicación $\cdot : K \times K \to K$ de K a $\Bbbk \times K$. De esta forma podemos ver al cuerpo $\mathbb{C}$ como un espacio vectorial sobre $\mathbb{R}$ o sobre $\mathbb{Q}$, por ejemplo. ◇

1.2.6. Proposición. *Sea $\Bbbk$ un cuerpo y sea V un espacio vectorial sobre $\Bbbk$.*

(*i*) *Para todo $x \in V$ es $0 \cdot x = 0_V$.*
(*ii*) *Para todo $a \in \Bbbk$ es $a \cdot 0_V = 0_V$.*
(*iii*) *Para todo $x \in V$ es $(-1) \cdot x = -x$.*

Demostración. Observemos primero que

$$\textit{si } x \in V \textit{ es tal que } x + x = x, \textit{ entonces } x = 0_V. \tag{1}$$

En efecto, si x tiene esa propiedad, tenemos que

$$x = x + 0_V = x + (x + (-x)) = (x + x) + (-x) = x + (-x) = 0_V.$$

Si $x \in V$, entonces

$$0 \cdot x + 0 \cdot x = (0 + 0) \cdot x = 0 \cdot x$$

y nuestra observación (1) nos dice que $0 \cdot x = 0_V$: esto prueba la afirmación (*i*) de la proposición. Si ahora $a \in \Bbbk$, entonces

$$a \cdot 0_V + a \cdot 0_V = a \cdot (0_V + 0_V) = a \cdot 0_V$$

y, usando otra vez la observación (1), concluimos que $a \cdot 0_V = 0_V$, como afirma la parte (*ii*) de la proposición. Para ver que vale la parte (*iii*), consideramos un elemento $x \in V$ y calculamos que

$$x + (-1) \cdot x = 1 \cdot x + (-1) \cdot x = (1 + (-1)) \cdot x = 0 \cdot x = 0_V$$

y, de manera similar, que $(-1) \cdot x + x = 0_V$. Estas dos igualdades nos dicen que $(-1) \cdot x$ es un elemento de V opuesto a x y, como sabemos que hay uno sólo, que $(-1) \cdot x = -x$. □

Convención

A partir de este momento fijaremos un cuerpo $\Bbbk$ arbitrario en todo lo que sigue y todos los espacios vectoriales que consideremos serán espacios vectoriales sobre $\Bbbk$. Por otro lado, escribiremos simplemente 0 para denotar tanto al cero del cuerpo $\Bbbk$ como al de todos los espacios vectoriales con los que estemos trabajando.

§3. Subespacios

1.3.1. Sea V un espacio vectorial. Un subconjunto U de V es un ***subespacio*** de V si

- **(S$_1$)** $0 \in U$;
- **(S$_2$)** si $x, y \in U$, entonces $x + y \in U$;
- **(S$_3$)** si $a \in \Bbbk$ y $x \in U$, entonces $a \cdot x \in U$;

Cuando ése es el caso, escribimos $U \leq V$.

1.3.2. Un subespacio de un espacio vectorial es, de manera natural, él mismo un espacio vectorial:

Proposición. *Sea V un espacio vectorial. Si U es un subespacio de V, entonces U es un espacio vectorial con respecto a las operaciones*

$$+ : (x, y) \in U \times U \longmapsto x + y \in U \qquad \cdot : (a, x) \in \Bbbk \times U \longmapsto a \cdot x \in U$$

obtenidas de las de V por restricción.

Notemos que es precisamente porque U satisface las condiciones (S$_2$) y (S$_3$) que las operaciones $+ : V \times V \to V$ y $\cdot : \Bbbk \times V \to V$ de V se restringen a operaciones $+ : U \times U \to U$ y $\cdot : \Bbbk \times U \to U$. Desde ahora en adelante, cada vez que consideremos un subespacio de un espacio vectorial, lo dotaremos implícitamente de la estructura de espacio vectorial que nos da esta proposición.

Demostración. Que las condiciones (V$_1$), (V$_2$), (V$_5$), (V$_6$) y (V$_7$) se cumplen en U es consecuencia inmediata de que se cumplen en V. La condición (S$_1$) implica que (V$_3$) se cumple en U, ya que 0, que está en U por hipótesis, es un elemento neutro para la suma de U. Resta ver, entonces, que la condición (V$_4$) se satisface en U. Pero sí $x \in U$, entonces sabemos que $x' = (-1) \cdot x \in U$ por (S$_3$) y que, según la Proposición 1.2.6*(iii)*, se tiene que $x + x' = 0$ y $x' + x = 0$ en V: en vista de la forma en que están definidas las operaciones de U, esto nos dice que x' es el opuesto de x. $\square$

1.3.3. Si V es un espacio vectorial, entonces los subconjuntos $\{0\}$ y V de V son subespacios de V —la verificación de las condiciones de la definición 1.3.1 es inmediata en los dos casos. Estos son los ***subespacios triviales*** de V. Casi siempre escribiremos simplemente 0 al subespacio $\{0\}$. Llamamos al subespacio 0 el ***subespacio nulo*** de V, llamamos al subespacio V de V el ***subespacio impropio*** de V y decimos que todos los demás son ***subespacios propios***.

1.3.4. Ejemplos.

(a) Sea $n \in \mathbb{N}$ y sea $V = \Bbbk^n$. Es fácil ver que los siguientes subconjuntos de V son subespacios de V:

$$U_1 = \{(x_1, \dots, x_n)^{\mathrm{t}} \in V : x_1 = 0\},$$
$$U_2 = \{(x_1, \dots, x_n)^{\mathrm{t}} \in V : x_1 + \cdots + x_n = 0\},$$
$$U_2 = \{(x_1, \dots, x_n)^{\mathrm{t}} \in V : x_1 = x_2, x_3 = x_4\}.$$

(b) Sea $X = (0, 1)$ el intervalo unidad abierto de $\mathbb{R}$ y sea $V = \mathbb{R}^X$ el espacio vectorial real de todas las funciones $X \to \mathbb{R}$, como en el Ejemplo 1.2.4(a). Los siguientes subconjuntos de V son subespacios de V:

$$U_1 = \{f \in V : f \text{ es continua}\},$$
$$U_2 = \{f \in V : f \text{ tiene derivada en todo punto de } X\},$$
$$U_3 = \{f \in V : f \text{ tiene derivada en todo punto de } X \text{ y } f' \text{ es continua en } X\},$$
$$U_4 = \{f \in V : f \text{ es continua y } f(\tfrac{1}{2}) = 0\},$$
$$U_5 = \{f \in V : f \text{ es continua y } \textstyle\int_{1/3}^{2/3} f(x)\,\mathrm{d}x = 0\}.$$
◇

1.3.5. Proposición. *Sea V un espacio vectorial.*

(*i*) *Si U_1 y U_2 son subespacios de V, entonces la intersección $U_1 \cap U_2$ también lo es.*

(*ii*) *Más generalmente, si $\{U_i : i \in I\}$ es una familia arbitraria de subespacios de V, entonces la intersección $\bigcap_{i\in I} U_i$ es un subespacio de V.*

Demostración. Basta que probemos la segunda parte, ya que contiene a la primera como un caso particular. Sea entonces $\{U_i : i \in I\}$ una familia de subespacios de V y pongamos $U = \bigcap_{i\in I} U_i$. Para ver que U es un subespacios de V, verificamos una a una las condiciones de la definición 1.3.1:

- Si $i \in I$, entonces $0 \in U_i$, porque U_i es un subespacio de V, y entonces $0 \in \bigcap_{i\in I} U_i = U$.
- Sean x e y dos elementos de U. Si $i \in I$, entonces x e y están en U_i, ya que $U \subseteq U_i$, y entonces $x + y \in U_i$, porque U_i satisface la condición (S_2). Esto nos dice que $x + y \in U$.
- Sean $a \in \Bbbk$ y $x \in U$. Si $i \in I$ entonces $x \in U_i$ y, como U_i es un subespacio de V, $a \cdot x \in U_i$. Vemos así que $a \cdot x \in \bigcap_{i\in I} U_i = U$.

Esto completa la prueba de la proposición. □

1.3.6. A diferencia de lo que ocurre con la intersección de subespacios de un espacio vectorial, la unión de subespacios no es en general un subespacio. De hecho, vale la siguiente observación:

Proposición. *Sea V un espacio vectorial y sean U_1 y U_2 dos subespacios de V. La unión $U_1 \cup U_2$ es un subespacio de V si y solamente si coincide con U_1 o con U_2.*

Notemos que esto último ocurre exactamente cuando o bien $U_1 \subseteq U_2$ o bien $U_2 \subseteq U_1$.

Demostración. Si $U_1 \cup U_2$ coincide con U_1 o con U_2, es claro que se trata de un subespacio. Veamos la implicación recíproca. Supongamos que $U_1 \cup U_2$ es un subespacio de V y que ni U_1 no está contenido en U_2 ni U_1 está contenido en U_2. Existen entonces vectores $x \in U_1 \setminus U_2$ e $y \in U_2 \setminus U_1$. Como ambos perteneces a $U_1 \cup U_2$ y esta unión es un subespacio de V, se sigue de esto que $x + y \in U_1 \cup U_2$. Esto es absurdo: si $x + y \in U_1$, entonces $y = (x + y) - y \in U_1$, contradiciendo la elección de y, y si $x + y \in U_2$, $x = (x + y) - y \in U_2$, contradiciendo la de x. □

§4. Combinaciones lineales

1.4.1. Sea V un espacio vectorial. Si $S \subseteq V$ es un subconjunto de V, decimos que un vector $v \in V$ es ***combinación lineal de los elementos de*** S si existen $n \in \mathbb{N}_0$, elementos $x_1, \dots, x_n$ de S y escalares $a_1, \dots, a_n \in \Bbbk$ tales que

$$x = a_1 x_1 + \cdots + a_n x_n.$$

En esta definición permitimos que n sea igual a 0: como una suma de cero elementos de V tiene valor 0, esto significa que el elemento 0 de V es una combinación lineal de los elementos de S. Escribimos $\langle S \rangle$ al conjunto de todos los elementos de V que son iguales a una combinación lineal de elementos de S. Cuando $S = \{x_1, \dots, x_n\}$, escribimos $\langle x_1, \dots, x_n \rangle$ en lugar de $\langle \{x_1, \dots, x_n\} \rangle$.

1.4.2. Proposición. *Sea V un espacio vectorial. Si S y T son subconjuntos de V y $S \subseteq T$, entonces se tiene que $\langle S \rangle \subseteq \langle T \rangle$.*

Demostración. En efecto, si $S \subseteq T \subseteq V$, entonces todo vector de V que es combinación lineal de elementos de S es, claramente, combinación lineal de elementos de T. □

1.4.3. Proposición. *Sea V un espacio vectorial. Si S es un subconjunto de V, entonces el subconjunto $\langle S \rangle$ es un subespacio de V que contiene a S.*

En vista de esto, tiene sentido que llamemos a $\langle S \rangle$ el ***subespacio generado por*** S.

Demostración. Sea $S \subseteq V$ un subconjunto. Como siempre, verificamos cada una de las condiciones de la definición 1.3.1:

- Que 0 es un elemento de $\langle S \rangle$ es consecuencia de la observación hecha en la definición 1.4.1.
- Supongamos que x e y son dos elementos de $\langle S \rangle$, de manera que existen $n, m \in \mathbb{N}_0$, elementos $x_1, \dots, x_n$ e $y_1, \dots, y_m$ de S y escalares $a_1, \dots, a_n, b_1, \dots, b_m \in \Bbbk$ tales que $x = a_1 x_1 + \cdots + a_n x_n$ e $y = b_1 y_1 + \cdots + b_m y_m$. Pongamos $z_i = x_i$ y $c_i = a_i$ para cada $i \in [\![n]\!]$ y $z_i = y_{i-n}$ y $c_i = b_{i-n}$ para cada $i \in \{n+1, \dots, n+m\}$. Tenemos entonces que

  $$x + y = a_1 x_1 + \cdots + a_n x_n + b_1 y_1 + \cdots + b_m y_m = c_1 z_1 + \cdots + c_{n+m} z_{n+m},$$

  y esto muestra que $x + y$ es una combinación lineal de elementos de S, esto es, que es un elemento de $\langle S \rangle$.
- Sean $a \in \Bbbk$ y $x \in \langle S \rangle$, de manera que existen $n \geq 0$, $x_1, \dots, x_n \in S$ y escalares $a_1, \dots, a_n \in \Bbbk$ tales que $x = a_1 x_1 + \cdots + a_n x_n$. Como

  $$a \cdot x = a \cdot (a_1 x_1 + \cdots + a_n x_n) = (a a_1) x_1 + \cdots + (a a_n) x_n,$$

  vemos que $a \cdot x$ es una combinación lineal de elementos de S y, entonces, que está en $\langle S \rangle$.

Para terminar, nos queda probar que $S \subseteq \langle S \rangle$, pero esto es inmediato: si $x \in S$, entonces $x = 1 \cdot x$ y esto muestra que x es una combinación lineal de elementos de S. □

1.4.4. Ejemplos.

(a) Sea sea $n \in \mathbb{N}$, sea $V = \Bbbk^n$ y para cada $i \in \llbracket n \rrbracket$ escribamos e_i al elemento

$$(0, \dots, 0, \underbrace{1}_{i}, 0, \dots, 0)$$

de $\Bbbk^n$ cuyas componentes son todas nulas salvo la i-ésima, que es igual a 1. El conjunto $S = \{e_1, \dots, e_n\}$ genera a V. En efecto, si $x = (x_1, \dots, x_n)$ es un vector de V, entonces claramente vale que

$$x = x_1 e_1 + \cdots + x_n e_n.$$

(b) Sea $\Bbbk[X]$ el espacio vectorial de los polinomios en la variable X con coeficientes en $\Bbbk$. El conjunto $S = \{X^i : i \in \mathbb{N}_0\}$ de los monomios genera a $\Bbbk[X]$ como espacio vectorial: en efecto, por la definición misma de lo que es un polinomio, todo elemento de $\Bbbk[X]$ es una combinación lineal de elementos de S.

(c) Sea X un conjunto y sea $V = \Bbbk^X$ el espacio vectorial de todas las funciones $X \to \Bbbk$. Para cada $x \in X$ sea $\delta_x : X \to \Bbbk$ la función tal que

$$\delta_x(y) = \begin{cases} 1, & \text{si } y = x; \\ 0, & \text{en cualquier otro caso.} \end{cases}$$

Sea, finalmente, $S = \{\delta_x : x \in X\}$. Afirmamos que

el conjunto S genera a V si y solamente si el conjunto X es finito.

Para ello, supongamos primero que X es finito, sea n el número de sus elementos y sean $x_1, \dots, x_n$ sus elementos. Si $f \in V$, entonces es inmediato verificar que

$$f = f(x_1)\delta_{x_1} + \cdots + f(x_n)\delta_{x_n}$$

y esto nos dice que $f \in \langle S \rangle$. Así, S genera a V en este caso.

Supongamos ahora que X es infinito y mostremos que S no genera a V. Sea $f : X \to \Bbbk$ el elemento de V tal que $f(x) = 1$ para cada $x \in X$ y, para llegar a un absurdo, supongamos que $f \in \langle S \rangle$, de manera que existen elementos $x_1, \dots, x_n \in X$ y escalares $a_1, \dots, a_n \in \Bbbk$ tales que

$$f = a_1 \delta_{x_1} + \cdots + a_n \delta_{x_n}. \tag{2}$$

Como estamos suponiendo que X es infinito, existe un elemento $x \in X$ distinto de todos los elementos $x_1, \dots, x_n$ y entonces evaluando ambos lados de la igualdad (2) en x vemos que

$$1 = f(x) = (a_1 \delta_{x_1} + \cdots + a_n \delta_{x_n})(x) = 0.$$

Esto, por supuesto, es absurdo, y podemos concluir que $V \supsetneq \langle S \rangle$.

(d) Si V es un espacio vectorial nulo, entonces $V = \langle \varnothing \rangle$. En efecto, sabemos que $\langle \varnothing \rangle$ es un subespacio de V, así que necesariamente contiene al elemento cero de V y, como ése es de hecho el único elemento de V, vale la igualdad que afirmamos. $\diamond$

1.4.5. Proposición. *Sea V un espacio vectorial, sea U un subespacio de V y sea S un subconjunto de V. El subespacio $\langle S\rangle$ está contenido en U si y solamente sí $S \subseteq U$.*

Demostración. Si $\langle S\rangle \subseteq U$, entonces $S \subseteq U$, ya que sabemos que $S \subseteq \langle S\rangle$: esto nos dice que la condición del enunciado es necesaria. Veamos que es suficiente.

Supongamos que $S \subseteq U$, sea $x \in \langle S\rangle$ y mostremos que $x \in U$. Como x es combinación lineal de elementos de S, existen $n \in \mathbb{N}_0$, $x_1, \dots, x_n \in S$ y $a_1, \dots, a_n \in \Bbbk$ tales que $x = a_1x_1 + \cdots + a_nx_n$. Para cada $i \in [\![n]\!]$ tenemos que $a_ix_i \in U$, ya que $x_i \in S \subseteq U$ y U es un subespacio, y entonces, juntando todo, vemos que $x = a_1x_1 + \cdots + a_nx_n \in U$, como queríamos. □

1.4.6. Proposición. *Sea V un espacio vectorial. Si S es un subconjunto de V, entonces el subespacio $\langle S\rangle$ es el menor subespacio de V que contiene a S y, de hecho, coincide con la intersección de todos los subespacios de V que contienen a S,*

$$\langle S\rangle = \bigcap_{\substack{U\leq V\\ S\subseteq U}} U.$$

Demostración. Llamemos W a la intersección que aparece en el enunciado. Como es una intersección de subespacios de V, la Proposición 1.3.5*(ii)* nos dice que W es un subespacio, y como todos los subespacios que aparecen en la intersección contienen a S, tenemos que $S \subseteq W$.

Afirmamos que W es el menor subespacio de V que contiene a S: en efecto, si U es otro subespacio de V que contiene a S, entonces U es uno de los espacios que aparecen en la intersección que define a W y, en consecuencia, es $W \subseteq U$.

En particular, como $\langle S\rangle$ es un subespacio de V que contiene a S, esto nos dice que $W \subseteq \langle S\rangle$. Por otro lado, como W contiene a S, la Proposición 1.4.5 nos dice que $\langle S\rangle \subseteq W$. Vemos así que $\langle S\rangle = W$ y esto completa la prueba de la proposición. □

§5. Dependencia e independencia lineal

1.5.1. Sea V un espacio vectorial sobre $\Bbbk$ y sea S un subconjunto de V. Decimos que S es ***linealmente dependiente*** si existen un entero positivo $n \in \mathbb{N}$, elementos $x_1, \dots, x_n$ de S distintos dos a dos y escalares $a_1, \dots, a_n \in \Bbbk$, no todos nulos, tales que $a_1x_1 + \cdots + a_nx_n = 0$, y decimos que es ***linealmente independiente*** en caso contrario. Observemos que el subconjunto vacío $\varnothing$ de V es linealmente independiente.

1.5.2. Proposición. *Sea V un espacio vectorial.*

(*i*) *Si $S \subseteq T \subseteq V$ y S es linealmente dependiente, entonces T es linealmente dependiente.*
(*ii*) *Si $S \subseteq T \subseteq V$ y T es linealmente independiente, entonces S es linealmente independiente.*
(*iii*) *Si $S \subseteq V$ y $0 \in S$, entonces S es linealmente dependiente.*

Demostración. (i) Sean S y T subconjuntos de V tales que $S \subseteq T$ y S es linealmente dependiente. Existen entonces $n \in \mathbb{N}$, $x_1, \dots, x_n \in S$ distintos dos a dos, y escalares $a_1, \dots, a_n \in \mathbb{k}$ no todos nulos tales que $a_1x_1 + \cdots + a_nx_n = 0$. Como $S \subseteq T$, cada uno de los vectores $x_1, \dots, x_n$ está en T, y es claro que esto implica que T es linealmente dependiente.

(ii) Esta afirmación es la contrarrecíproca de la de (i).

(iii) Si $S \subseteq V$ contiene a 0, podemos tomar $n = 1$, $x_1 = 0$ y $a_1 = 1$: como $a_1x_1 = 0$ en ese caso, el conjunto S es linealmente dependiente. □

1.5.3. Proposición. *Sea V un espacio vectorial. Un subconjunto S de V es linealmente independiente si y solamente si satisface la siguiente condición:*

> *cada vez que $x_1, \dots, x_n$ son elementos de S distintos dos a dos y $a_1, \dots, a_n \in \mathbb{k}$ son escalares tales que $a_1x_1 + \cdots + a_nx_n = 0$, se tiene que de hecho $a_1 = \cdots = a_n = 0$.*

Demostración. En efecto, la condición que aparece en el enunciado dice simplemente que el conjunto no es linealmente dependiente. □

1.5.4. Proposición. *Sea V un espacio vectorial y sea $S \subseteq V$ un subconjunto. Las siguientes dos afirmaciones son equivalentes:*

(a) *El conjunto S es linealmente dependiente.*
(b) *Existe $x \in S$ tal que $x \in \langle S \setminus \{x\} \rangle$ y $\langle S \setminus \{x\} \rangle = \langle S \rangle$.*

Demostración. (a) ⇒ (b) Supongamos que S es linealmente dependiente, de manera que existen un entero positivo $n \in \mathbb{N}$, vectores $x_1, \dots, x_n \in S$ distintos dos a dos y escalares $a_1, \dots, a_n \in \mathbb{k}$ no todos nulos, tales que

$$a_1x_1 + \cdots + a_nx_n = 0. \tag{3}$$

Como los escalares no son todos nulos, existe $i \in [\![n]\!]$ tal que $a_i \neq 0$ y entonces de la igualdad (3) vemos que

$$x_i = (-a_i^{-1}a_1)x_1 + \cdots + \widehat{(-a_i^{-1}a_i)x_i} + \cdots + (-a_i^{-1}a_n)x_n.$$

Los vectores $x_1, \dots, \widehat{x_i}, \dots, x_n$ que aparecen aquí a la derecha están todos en $S \setminus \{x_i\}$, así que $x_i \in \langle S \setminus \{x_i\} \rangle$. Esto prueba lo que queremos.

(b) ⇒ (a) Supongamos que existe $x \in S$ tal que $x \in \langle S \setminus \{x\} \rangle$, de manera que existen $n \in \mathbb{N}_0$, $x_1, \dots, x_n \in S$ y $a_1, \dots, a_n \in \mathbb{k}$ tales que

$$x = a_1x_1 + \cdots + a_nx_n. \tag{4}$$

De entre todas las elecciones posibles de n, de los vectores $x_1, \dots, x_n$ y de los escalares $a_1, \dots, a_n$, que generalmente son muchas, quedémonos que alguna que tiene *n mínimo*. En ese caso, los vectores $x_1, \dots, x_n$ son distintos dos a dos: en efecto, si no fuese el caso y tuviéramos, por ejemplo, que x_1 y x_2 son iguales, tendríamos que $x = (a_1 + a_2)x_2 + a_3x_3 + \cdots + a_nx_n$, contradiciendo la minimalidad de n. Por otro lado, como esos n vectores están en $S \setminus \{x\}$, son todos distintos de x y, en consecuencia, vemos que los $n+1$ vectores $x, x_1, \ldots, x_n$ son distintos dos a dos.

De la igualdad (4) se deduce que

$$(-1)x + a_1x_1 + \cdots + a_nx_n = 0$$

y, como los $n+1$ vectores que aparecen aquí están en S y son distintos dos a dos y claramente no todos los coeficientes son nulos, esto muestra que S es linealmente dependiente. □

1.5.5. Lema. *Sea V un espacio vectorial, sea S un subconjunto de V linealmente independiente y sea $x \in V \setminus S$. El conjunto $S \cup \{x\}$ es linealmente independiente si y solamente si $x \in V \setminus \langle S \rangle$.*

Demostración. Supongamos primero que el conjunto $S \cup \{x\}$ es linealmente dependiente, de manera que existen $n \in \mathbb{N}$, elementos $x_1, \ldots, x_n \in S \cup \{x\}$ distintos dos ados y escalares $a_1, \ldots, a_n \in \Bbbk$ no nulos tales que

$$a_1x_1 + \cdots + a_nx_n = 0. \tag{5}$$

No puede ser que todos los vectores $x_1, \ldots, x_n$ sean distintos de x: en ese caso estarían todos en S y esto es imposible ya que S es linealmente independiente. Podemos suponer, entonces, que $x_1 = x$ y de (5) deducir que

$$x = x_1 = \left(-\frac{a_2}{a_1}\right) + \cdots + \left(-\frac{a_2}{a_1}\right) \in \langle S \rangle.$$

Para probar la implicación recíproca, supongamos ahora que $x \in \langle S \rangle$, de manera que existen $n \in \mathbb{N}_0$, elementos $x_1, \ldots, x_n \in S$ distintos dos a dos y escalares $a_1, \ldots, a_n \in \Bbbk$ tales que $x = a_1x_1 + \cdots + a_nx_n$. Como x no pertenece a S por hipótesis, los $n+1$ vectores $x, x_1, \ldots, x_n$ de $S \cup \{x\}$ son distintos dos a dos y, como claramente $(-1)x + a_1x_1 + \cdots + a_nx_n = 0$, el conjunto $S \cup \{x\}$ es linealmente dependiente. □

§6. Bases

1.6.1. Sea V un espacio vectorial. Decimos que un subconjunto $\mathscr{B}$ de V es una ***base*** de V si

- $\mathscr{B}$ es un conjunto linealmente independiente, y
- $\mathscr{B}$ genera a V, de manera que $V = \langle \mathscr{B} \rangle$.

1.6.2. La siguiente proposición explicita las propiedades más importantes que tienen las bases de un espacio vectorial:

Proposición. *Sea V un espacio vectorial. Si $\mathscr{B}$ es un subconjunto de V, entonces las siguientes afirmaciones son equivalentes:*

(a) *$\mathscr{B}$ es una base de V.*

(b) *$\mathscr{B}$ es un conjunto linealmente independiente y es maximal con esta propiedad: todo subconjunto S de V tal que $\mathscr{B} \subsetneq S$ es linealmente dependiente.*

(c) *$\mathscr{B}$ genera a V y es minimal con esta propiedad: ningún subconjunto propio de $\mathscr{B}$ genera a V.*

Demostración. Sea $\mathscr{B}$ un subconjunto de V.

$(a) \Rightarrow (b)$ Supongamos que $\mathscr{B}$ es una base de V. En vista de la definición 1.6.1, el conjunto $\mathscr{B}$ es linealmente independiente. Mostremos que es maximal con esta propiedad.

Sea $S \subseteq V$ un subconjunto tal que $S \supsetneq \mathscr{B}$, de manera que existe algún vector $x \in S \setminus \mathscr{B}$. Como $\mathscr{B}$ es una base de V, genera a V y entonces $x \in \langle \mathscr{B} \rangle$. Como $x \in S \setminus \mathscr{B}$, es claro que $\mathscr{B} \subseteq S \setminus \{x\}$ y, en consecuencia, que $x \in \langle \mathscr{B} \rangle \subseteq \langle S \setminus \{x\} \rangle$. La Proposición 1.5.4 nos dice, entonces, que S es linealmente dependiente.

$(b) \Rightarrow (c)$ Sea $\mathscr{B}$ un subconjunto de V linealmente independiente y maximal con esta propiedad. Veamos que $\mathscr{B}$ genera a V y que es minimal con esta propiedad.

Sea $x \in V$. Si $x \in \mathscr{B}$, es claro que $x \in \langle \mathscr{B} \rangle$, así que supongamos que no es ése el caso. Entonces $\mathscr{B} \cup \{x\}$ es un subconjunto de V que contiene propiamente a $\mathscr{B}$ y la maximalidad de este último implica que $\mathscr{B} \cup \{x\}$ es un conjunto linealmente dependiente. Existen entonces $n \in \mathbb{N}$, vectores $x_1, \dots, x_n \in \mathscr{B} \cup \{x\}$ distintos dos a dos y escalares $a_1, \dots, a_n \in \Bbbk$ no nulos tales que

$$a_1 x_1 + \cdots + a_n x_n = 0. \tag{6}$$

Observemos que no puede ser que todos los vectores $x_1, \dots, x_n$ pertenezcan a $\mathscr{B}$, porque este conjunto es linealmente independiente: uno de ellos es entonces igual a x y, sin pérdida de generalidad, podemos suponer que, de hecho, es $x_1 = x$ y, por lo tanto, que los otros vectores $x_2, \dots, x_n$ pertenecen a $\mathscr{B}$. En ese caso, la igualdad (6) implica que

$$x = x_1 = a_1^{-1} a_2 x_2 + \cdots + a_1^{-1} a_n x_n \in \langle \mathscr{B} \rangle.$$

Esto muestra que $\langle \mathscr{B} \rangle = V$, esto es, que $\mathscr{B}$ genera a V, como queríamos.

Sea ahora T un subconjunto propio de $\mathscr{B}$ y supongamos que T genera a V. Como $T \subsetneq \mathscr{B}$, existe un vector $x \in \mathscr{B} \setminus T$, y como $x \in V = \langle T \rangle$, existen $n \in \mathbb{N}_0$, vectores $x_1, \dots, x_n \in T$ distintos dos a dos y escalares $a_1, \dots, a_n \in \Bbbk$ tales que $x = a_1 x_1 + \cdots + a_n x_n$. Esto nos dice que

$$a_1 x_1 + \cdots + a_n x_n + (-1)x = 0. \tag{7}$$

Como los vectores $x_1, \dots, x_n$ son distintos dos a dos y son todos distintos de x, ya que están en T y x no, y los escalares que aparecen en la combinación lineal (7) no son todos nulo, vemos así que el conjunto $\mathscr{B}$ es linealmente dependiente. Esto es absurdo, ya que contradice nuestra hipótesis sobre $\mathscr{B}$. Esta contradicción provino de suponer que T, que es un subconjunto propio de $\mathscr{B}$, genera a V, y prueba, entonces, que $\mathscr{B}$ es un subconjunto generador minimal de V, como queríamos.

$(c) \Rightarrow (a)$ Sea $\mathscr{B}$ un subconjunto de V que genera a V y que es minimal con esa propiedad. Para ver que se trata de una base, alcanza con que mostremos que es linealmente independiente.

Supongamos, para llegar a una contradicción, que no lo es. La Proposición 1.5.4 nos dice entonces que existe $x \in \mathscr{B}$ tal que $\langle \mathscr{B} \setminus \{x\} \rangle = \langle \mathscr{B} \rangle$ y esto es absurdo, ya que en ese caso $S \setminus \{x\}$ es un subconjunto propio de $\mathscr{B}$ que genera a V. Esto completa la prueba de la proposición. □

1.6.3. Decimos que un espacio vectorial V es ***finitamente generado*** si existe un subconjunto *finito* $S \subseteq V$ que genera a V, esto es, tal que $V = \langle S \rangle$. Si V satisface esta condición podemos mostrar que V posee una base:

Proposición. *Sea V un espacio vectorial. Si V es finitamente generado, entonces V posee una base finita. Más precisamente, si S es un subconjunto finito de V que lo genera, entonces existe una base de V contenida en S.*

Demostración. Sea $S = \{x_1, \dots, x_n\}$ un subconjunto finito de V tal que $\langle S \rangle = V$. Claramente S posee subconjuntos que generan a V —por ejemplo, S mismo es uno de ellos— y entonces hay un subconjunto T de S de cardinal mínimo entre todos ellos: esto es, existe un subconjunto T de S que genera a V y tal que todo subconjunto de S con menos elementos que T no genera a V. En particular, ningún subconjunto *propio* de T genera a V y, de acuerdo a la Proposición 1.6.2, esto significa que T es una base de V. □

§7. Dimensión

1.7.1. Proposición. *Sea V un espacio vectorial y sea $n \in \mathbb{N}_0$. Si V posee una base que tiene n elementos, entonces todo subconjunto de V con* más *que n elementos es linealmente dependiente.*

Demostración. Supongamos que $\mathscr{B} = \{x_1, \dots, x_n\}$ es una base de V, sea $T = \{y_1, \dots, y_m\}$ un conjunto linealmente independiente de m elementos y supongamos, para llegar a un absurdo, que $m > n$. Mostremos que para cada $r \in \{0, \dots, n\}$ vale que

existen índices $i_1, \dots, i_{n-r} \in [\![n]\!]$, distintos dos a dos, tales que el conjunto $\{y_1, \dots, y_r, x_{i_1}, \dots, x_{i_{n-r}}\}$ es una base de V. (8)

Para hacerlo, procedemos por inducción en r. Notemos que cuando $r = 0$ no hay nada que probar, ya que basta tomar $i_j = j$ para cada $j \in [\![n]\!]$.

Sea entonces $r \in \{0, \dots, n-1\}$ y supongamos que existen índices $i_1, \dots, i_{n-r} \in [\![n]\!]$, distintos dos a dos, tales que el conjunto $S = \{y_1, \dots, y_r, x_{i_1}, \dots, x_{i_{n-r}}\}$ es una base de V. Como este conjunto es una base, sabemos que existen escalares $a_1, \dots, a_r, b_1, \dots, b_{n-r} \in \Bbbk$ tales que

$$y_{r+1} = a_1 y_1 + \cdots + a_r y_r + b_1 x_{i_1} + \cdots + b_{n-r} x_{i_{n-r}}. \tag{9}$$

No puede ser que todos los escalares $b_1, \dots, b_{n-r}$ sean nulos: de ser ése el caso, a partir de la igualdad (9) obtendríamos una combinación lineal nula no trivial de elementos del conjunto T, que es linealmente independiente. Así, alguno de esos escalares es no nulo y, sin pérdida de generalidad, podemos suponer que se trata del último, b_{n-r}. De la igualdad (9), entonces, podemos deducir que

$$x_{i_{n-r}} = \Big(-\frac{a_1}{b_{n-r}}\Big)y_1 + \cdots + \Big(-\frac{a_r}{b_{n-r}}\Big)y_r + \frac{1}{b_{n-r}}y_{r+1} + \Big(-\frac{b_1}{b_{n-r}}\Big)x_{i_1} + \cdots + \Big(-\frac{b_{n-r-1}}{b_{n-r}}\Big)x_{i_{n-r-1}}.$$

Sea $S' = \{y_1, \dots, y_{r+1}, x_{i_1}, \dots, x_{i_{n-r-1}}\}$. La última igualdad nos dice que $x_{n-r} \in \langle S' \rangle$ y entonces que

$$\langle S' \rangle = \langle S' \cup \{x_{n-r}\} \rangle = \langle S \cup \{y_{r+1}\} \rangle \supseteq \langle S \rangle = V,$$

de manera que S' genera a V. Para ver que S' es una base nos resta probar que es linealmente independiente. Sean entonces $c_1, \dots, c_{r+1}, d_1, \dots, d_{n-r-1} \in \Bbbk$ escalares tales que

$$c_1 y_1 + \cdots + c_{r+1} y_{r+1} + d_1 x_{i_1} + \cdots + d_{n-r-1} x_{i_{n-r-1}} = 0. \tag{10}$$

Ahora bien, si en esta igualdad usamos la igualdad (9) para eliminar y_{r+1}, vemos que

$$\begin{aligned} c_1 y_1 + \cdots + c_r y_r + c_{r+1}\big(a_1 y_1 + \cdots + a_r y_r + b_1 x_{i_1} + \cdots + b_{n-r} x_{i_{n-r}}\big) \\ + d_1 x_{i_1} + \cdots + d_{n-r-1} x_{i_{n-r-1}} = 0 \end{aligned}$$

y, agrupando términos, que

$$\begin{aligned} (c_1 + c_{r+1}a_1)y_1 + \cdots + (c_r + c_{r+1}a_r)y_r + (c_{r+1}b_1 + d_1)x_{i_1} + \cdots + (c_{r-1}b_{n-r-1} + d_{n-r-1})x_{i_{n-r-1}} \\ + c_{r+1}b_{n_r}x_{i_{n-r}} = 0. \end{aligned}$$

Como el conjunto S es linealmente independiente, esto nos dice que los coeficientes que aparecen en esta combinación lineal son todos nulos, esto es, que valen las igualdades

$$\begin{aligned} &c_1 + c_{r+1}a_1 = \cdots = c_r + c_{r+1}a_r = 0, \\ &c_{r+1}b_1 + d_1 = \cdots = c_{r-1}b_{n-r-1} + d_{n-r-1} = 0, \\ &c_{r+1}b_{n-r} = 0. \end{aligned}$$

Recordando que el escalar b_{n-r} no es nulo, la última de estas ecuaciones nos dice que $c_{r+1} = 0$ y, usando eso, todas las otras nos dicen que $c_1 = \cdots = c_r = 0$ y que $d_1 = \cdots = d_{n-r-1} = 0$. Así, hemos concluido que todos los coeficientes que aparecen en la combinación lineal (10) son nulos: esto significa que el conjunto S' es linealmente independiente, como queríamos probar.

Esto completa la inducción que prueba nuestra afirmación (8). En particular, si tomamos $r = n$ ahí, vemos que el conjunto $\{y_1, \dots, y_n\}$ es una base de V. Ahora bien, esto significa, entre otras cosas, que existen escalares $e_1, \dots, e_n \in \Bbbk$ tales que $y_{n+1} = e_1 y_1 + \cdots + e_n y_n$, y esto es absurdo ya que el conjunto T es linealmente independiente. Esta contradicción muestra que nuestra hipótesis de que m es estrictamente más grande que n no puede satisfacerse y, en consecuencia, la proposición. □

1.7.2. La afirmación (8) en la que basamos la prueba de esta proposición es conocida como el Lema de Intercambio de Steinitz, por Ernst Steinitz (1871–1928).

1.7.3. La consecuencia más importante de la Proposición 1.7.1 es el siguiente resultado, que es fundamental en todo lo que sigue:

Teorema. *Sea V un espacio vectorial. Si V posee alguna base finita, entonces todas sus bases son finitas y tienen el mismo número de elementos.*

Demostración. Supongamos que $\mathscr{B} = \{x_1, \dots, x_n\}$ es una base finita de V con n elementos y que $\mathscr{B}'$ es otra base cualquiera de V. Si $\mathscr{B}'$ tiene *más* que n elementos, podemos elegir $n+1$ vectores $y_1, \dots, y_{n+1}$ distintos dos a dos en $\mathscr{B}'$ y, en consecuencia, el conjunto $\{y_1, \dots, y_{n+1}\}$ es linealmente independiente: esto es absurdo, ya que contradice a la Proposición 1.7.1. Vemos así que el conjunto $\mathscr{B}'$ es necesariamente finito y que tiene a lo sumo n elementos. Esto prueba la primera afirmación del teorema. Intercambiando los roles de $\mathscr{B}$ y de $\mathscr{B}'$ en el razonamiento que acabamos de hacer, vemos que también prueba la segunda afirmación. □

1.7.4. Decimos que un espacio vectorial V tiene ***dimensión finita*** si posee una base finita y, en ese caso, que el cardinal de esa base es la ***dimensión*** de V, que escribimos $\dim V$; en caso contrario decimos que V tiene ***dimensión infinita***. El Teorema 1.7.3 nos dice que todas las bases de un espacio vectorial de dimensión finita tienen la misma cantidad de elementos, así que la definición de la dimensión tiene sentido.

1.7.5. Proposición. *Sea V un espacio vectorial. Si V tiene dimensión infinita, entonces existe un subconjunto $S \subseteq V$ infinito y linealmente independiente.*

Demostración. Supongamos que V no tiene dimensión finita y mostremos inductivamente que

> *existe una sucesión $(S_i)_{i\geq 0}$ de subconjuntos de V linealmente independientes y tales que para cada $i \in \mathbb{N}_0$ el conjunto S_i tiene exactamente i elementos y $S_i \subsetneq S_{i+1}$.* (11)

Empezamos poniendo $S_0 = \varnothing$, que es un subconjunto finito linealmente independiente de V con 0 elementos. Para continuar, supongamos que $n \in \mathbb{N}_0$ y que ya construimos el conjunto S_n. Como S_n es finito, no puede ser que $\langle S_n \rangle$ sea igual a V, porque en ese caso S_n contendría, de acuerdo a la Proposición 1.6.3, una base de V y ésta sería finita, contradiciendo nuestra hipótesis sobre V. Así, existe un vector $x \in V \setminus \langle S_n \rangle$ y podemos considerar el conjunto $S_{n+1} = S_n \cup \{x\}$, que claramente contiene a S_n. Como $x \notin \langle S_n \rangle$, en particular $x \notin S_n$ y, en consecuencia, el conjunto S_{n+1} tiene $n+1$ elementos. Veamos que S_{n+1} es linealmente independiente. Para ello, supongamos que, por el contrario, existen elementos $x_1, \dots, x_k$ de S_{n+1} distintos dos a dos y escalares $a_1, \dots, a_k \in \Bbbk$ no nulos tales que

$$a_1x_1 + \cdots + a_kx_n = 0. \tag{12}$$

Los vectores $x_1, \dots, x_k$ no pueden ser todos distintos de x, porque en ese caso pertenecerían todos a S_n y la igualdad (12) nos diría que S_n es linealmente independiente. Sin pérdida de generalidad, entonces, podemos suponer que $x_1 = x$ y, en consecuencia, que $x_2, \dots, x_k \in S_n$. Se sigue entonces de (12) que

$$x = x_1 = \Big(-\frac{a_2}{a_1}\Big)x_2 + \cdots + \Big(-\frac{a_k}{a_1}\Big)x_k \in \langle S_n \rangle,$$

y esto es imposible en vista de la forma en que elegimos el vector x. Esta contradicción nos dice que no puede ser que S_{n+1} sea linealmente dependiente, así que es linealmente independiente.

Todo esto prueba la afirmación (11). Para probar la proposición, consideremos el conjunto $S = \bigcup_{n\geq 1} S_n$, que es infinito, y mostremos que es linealmente independiente.

Si no lo fuese, existirían elementos $y_1, \ldots, y_k \in S$ distintos dos a dos y escalares $b_1, \ldots, b_k \in \Bbbk$ no nulos tales que

$$b_1 y_1 + \cdots + b_k y_k = 0. \tag{13}$$

Ahora bien, para cada $i \in [\![k]\!]$ existe $m_i \in \mathbb{N}_0$ tal que $y_i \in S_{m_i}$, y podemos considerar el entero $m = \max\{m_1, \ldots, m_k\}$. Se sigue de (11) que $y_1, \ldots, y_k \in S_m$ y entonces la igualdad (13) implica que el conjunto S_m es linealmente dependiente. Esto es absurdo. □

1.7.6. Proposición. *Sea V un espacio vectorial de dimensión finita n. Si U es un subespacio de V, entonces U tiene dimensión finita y vale que* $\dim U \leq \dim V$.

Demostración. Sea U un subespacio de V y supongamos que no tiene dimensión finita. La Proposición 1.7.5 nos dice que entonces existe un subconjunto $S \subseteq U$ infinito y linealmente independiente. Como S está contenido en V, esto es absurdo en vista de la Proposición 1.7.1.

Vemos así que U tiene que ser necesariamente de dimensión finita. Sea $\mathscr{B}$ una base de U. Como $\mathscr{B}$ es un subconjunto linealmente independiente de V, la Proposición 1.7.1 afirma que tiene a lo sumo n elementos, y esto significa precisamente que $\dim U \leq \dim V$. □

1.7.7. Proposición. *Sea V un espacio vectorial de dimensión finita n y consideremos un subconjunto $S = \{x_1, \ldots, x_n\}$ de V de exactamente n elementos. Si S es linealmente independiente o si S genera a V, entonces S es una base de V.*

Demostración. Supongamos primero que S es linealmente independiente. Si no es una base, la Proposición 1.6.2 nos dice que existe un subconjunto T de V que contiene estrictamente a S y que es linealmente independiente y, en vista de la Proposición 1.7.1, esto es imposible, ya que este conjunto T tiene más que n elementos.

Supongamos ahora que S genera a V. Si S no es una base, la Proposición 1.6.2 implica que existe un subconjunto propio T de S que genera a V y, según Proposición 1.6.3, este conjunto T contiene una base $\mathscr{B}$ de V. Claramente, esta base $\mathscr{B}$ tiene *menos* elementos que S, y esto es absurdo, ya que $\dim V = n$. Así, el conjunto S debe ser una base, como queríamos. □

1.7.8. Proposición. *Sea V un espacio vectorial de dimensión finita n. Si $0 \leq r < n$ y $\{x_1, \ldots, x_r\}$ es un conjunto con r elementos que es linealmente independiente, entonces existen vectores $x_{r+1}, \ldots, x_n$ en V tales que $\{x_1, \ldots, x_n\}$ es una base de V.*

Demostración. Sea $r \in \{0, \dots, n-1\}$ y sea $S = \{x_1, \dots, x_r\}$ un subconjunto linealmente independiente de V de r elementos. Claramente existen subconjuntos linealmente independientes de V que contienen a S —por ejemplo, el conjunto S mismo— y todos ellos tienen a lo sumo n elementos, de acuerdo a la Proposición 1.7.1. Podemos entonces considerar el número $m \in \mathbb{N}_0$ máximo con la propiedad de que hay en V un subconjunto linealmente independiente T con m elementos y tal que $S \subseteq T$. Por supuesto, es $m \leq n$.

Fijemos un tal subconjunto T linealmente independiente de V con m elementos y que contiene a S, y supongamos que es $m < n$. El conjunto T no puede generar a V: si lo generase, sería una base de V con menos de n elementos, lo que es imposible. Existe entonces un vector $x \in V \setminus \langle T \rangle$ y, de acuerdo al Lema 1.5.5, el conjunto $T' = T \cup \{x\}$ es linealmente independiente. Esto es absurdo, ya que T' contiene a S y tiene *más* que m elementos, contradiciendo la forma en que elegimos al número m. Así, no puede ser que m sea menor que n y, en consecuencia, es $m = n$.

Basta entonces que elijamos como $x_{r+1}, \dots, x_n$ a los $n - r$ elementos del conjunto $T \setminus S$ para obtener la conclusión del enunciado. □

1.7.9. Proposición. *Sea V un espacio vectorial de dimensión finita y sea U un subespacio de V.*

(*i*) *Es $U = V$ si y solamente si $\dim U = \dim V$.*

(*ii*) *Es $U = 0$ si y solamente si $\dim U = 0$.*

Demostración. (*i*) La condición es evidentemente necesaria. Para ver que es suficiente, supongamos que $\dim U = \dim V$ y sea $\mathscr{B}$ una base de U. Se tiene entonces, por supuesto, que $\langle \mathscr{B} \rangle = U$; por otro lado, como $\mathscr{B}$ es un subconjunto linealmente independiente de V con cantidad de elementos igual a $\dim V$, genera a V y, en consecuencia, $\langle \mathscr{B} \rangle = V$. Así, es $U = V$, como queremos.

(*ii*) Sabemos que la dimensión de un espacio nulo es 0. Recíprocamente, si $\dim U = 0$, entonces el conjunto vacío $\varnothing$ genera a U y esto es posible solamente si $U = 0$. □

1.7.10. Proposición. *Sea V un espacio vectorial de dimensión finita. Si $(F_r)_{r \in \mathbb{N}}$ es una sucesión de subespacios de V tales que*

- *o bien $F_r \subseteq F_{r+1}$ para todo $r \in \mathbb{N}$*
- *o bien $F_r \supseteq F_{r+1}$ para todo $r \in \mathbb{N}$,*

entonces existe $r_0 \in \mathbb{N}$ tal que $F_r = F_{r_0}$ para todo $r \geq r_0$.

Demostración. **Hacer.** □

§8. Coordenadas y cambio de base

1.8.1. Sea V un espacio vectorial de dimensión finita n y sea $\mathscr{B} = (x_1, \dots, x_n)$ una base ordenada de V. Para cada $x \in V$ existen escalares $a_1, \dots, a_n \in \Bbbk$ tales que $x = a_1 x_1 + \cdots + a_n x_n$, y estos escalares están unívocamente determinados por x: llamamos al vector $(a_1, \dots, a_n)^{\mathsf{t}} \in \Bbbk^n$ el ***vector de coordenadas*** de X con respecto a la base $\mathscr{B}$ y lo denotamos $[x]_{\mathscr{B}}$.

La observación más importante sobre esta construcción es la siguiente:

Proposición. *Sea V un espacio vectorial de dimensión finita n y sea $\mathscr{B}$ una base ordenada de V. La función $c_{\mathscr{B}} : x \in V \mapsto [x]_{\mathscr{B}} \in \Bbbk^n$ es una biyección.*

Demostración. Consideremos la función $d : \Bbbk^n \to V$ tal que para cada $a = (a_1, \dots, a_n) \in \Bbbk^n$ es $d(a) = a_1 x_1 + \cdots + a_n x_n$. Para probar la proposición es suficiente que mostremos que $c_{\mathscr{B}}$ y d son funciones mutuamente inversas.

- Si $x \in V$ y $c_{\mathscr{B}}(x) = (a_1, \dots, a_n)$, entonces de la definición misma de la función $c_{\mathscr{B}}$ se sigue que $d(c_{\mathscr{B}}(x)) = a_1 x_1 + \cdots + a_n x_n = x$.
- Por otro lado, si $a = (a_1, \dots, a_n) \in \Bbbk^n$ y x es el vector $d(a) = a_1 x_1 + \cdots + a_n x_n$ de V, entonces claramente $c_{\mathscr{B}}(d(a)) = c_{\mathscr{B}}(x) = a$.

Vemos asi que $d \circ c_{\mathscr{B}} = \mathrm{id}_V$ y que $c_{\mathscr{B}} \circ d = \mathrm{id}_{\Bbbk^n}$, como queríamos. □

1.8.2. Sea V un espacio vectorial de dimensión finita n y sean $\mathscr{B} = (x_1, \dots, x_n)$ y $\mathscr{B}' = (y_1, \dots, y_n)$ dos bases ordenadas de V. Para cada $i \in [\![n]\!]$ sabemos que existen escalares $c_{1,i}, \dots, c_{n,i} \in \Bbbk$ tales que $x_i = c_{1,i} y_1 + \cdots + c_{n,i} y_1$, y podemos entonces considerar la matriz

$$C(\mathscr{B}, \mathscr{B}') = \begin{pmatrix} c_{1,1} & \cdots & c_{1,n} \\ \vdots & \ddots & \vdots \\ c_{n,1} & \cdots & c_{n,n} \end{pmatrix} \in \mathsf{M}_n(\Bbbk),$$

a la que llamamos ***matriz de cambio de base de $\mathscr{B}$ a $\mathscr{B}'$***: sus columnas de esta matriz son lso vectores de coordenadas de los elementos de la base $\mathscr{B}$ con respecto a la base $\mathscr{B}'$.

1.8.3. La razón por la que nos interesamos en esta matriz es que permite expresar las coordenadas de un vector cualquiera de un espacio vectorial con respecto de una base en términos de sus coordenadas con respecto a otra:

Proposición. *Sea V un espacio vectorial de dimensión finita y sean $\mathscr{B}$ y $\mathscr{B}'$ dos bases ordenadas de V. Para cada $x \in V$ se tiene que*

$$[x]_{\mathscr{B}'} = C(\mathscr{B}, \mathscr{B}') \cdot [x]_{\mathscr{B}}$$

y, de hecho, la matriz $C(\mathscr{B}, \mathscr{B}')$ es la única con esta propiedad.

Demostración. Supongamos que $\mathscr{B} = (x_1, \dots, x_n)$, $\mathscr{B}' = (y_1, \dots, y_n)$ y $C(\mathscr{B}, \mathscr{B}') = (c_{i,j})$, de manera que para cada $i \in [\![n]\!]$ se tiene que

$$x_i = c_{1,i} y_1 + \cdots + c_{n,i} y_i.$$

y sea $x \in V$ y $[x]_{\mathscr{B}} = (a_1, \dots, a_n)$. Tenemos entonces que

$$\begin{aligned} x &= a_1 x_1 + \cdots + a_n x_n \\ &= a_1(c_{1,1} y_1 + \cdots + c_{n,1} y_n) + \cdots + a_n(c_{1,n} y_1 + \cdots + c_{n,n} y_n) \\ &= (c_{1,1} a_1 + \cdots + c_{1,n} a_n) y_1 + \cdots + (c_{n,1} a_1 + \cdots + c_{n,n} a_n) y_n \end{aligned}$$

y esto significa que para cada $i \in [\![n]\!]$ la coordenada i-ésima de x con respecto a la base $\mathscr{B}'$ es precisamente la componente i-ésima del vector $C(\mathscr{B}, \mathscr{B}') \cdot [x]_{\mathscr{B}}$. Esto prueba la primera afirmación de la proposición.

Para ver la segunda, basta observar que si $A \in \mathsf{M}_n(\Bbbk)$ es una matriz tal que $[x]_{\mathscr{B}'} = A \cdot [x]_{\mathscr{B}}$ para todo $x \in V$, en particular se tiene que —como el vector de coordenadas $[x_i]_{\mathscr{B}}$ es el i-ésimo vector e_i de la base ordenada estándar de $\Bbbk^n$— la columna i-ésima de A es

$$A \cdot e_i = A \cdot [x_i]_{\mathscr{B}} = [x_i]_{\mathscr{B}'},$$

Vemos así que la matriz A es entonces la matriz $C(\mathscr{B}, \mathscr{B}')$. □

1.8.4. Proposición. *Sea V un espacio vectorial de dimensión finita n.*

(*i*) *Si $\mathscr{B}$ es una base ordenada de V, entonces $C(\mathscr{B}, \mathscr{B}) = I_n$, la matriz identidad de $\mathsf{M}_n(\Bbbk)$.*

(*ii*) *Si $\mathscr{B}$, $\mathscr{B}'$ y $\mathscr{B}''$ son bases ordenadas de V, entonces*

$$C(\mathscr{B}, \mathscr{B}'') = C(\mathscr{B}', \mathscr{B}'') \cdot C(\mathscr{B}, \mathscr{B}').$$

(*iii*) *Si $\mathscr{B}$ y $\mathscr{B}'$ son bases ordenadas de V, entonces la matriz $C(\mathscr{B}, \mathscr{B}')$ es inversible y*

$$C(\mathscr{B}, \mathscr{B}')^{-1} = C(\mathscr{B}', \mathscr{B}).$$

Demostración. La primera afirmación es evidente. Veamos la segunda: sean $\mathscr{B} = (x_1, \dots, x_n)$, $\mathscr{B}' = (y_1, \dots, y_n)$ y $\mathscr{B}'' = (z_1, \dots, z_n)$ tres bases ordenandas de V, y sean $C(\mathscr{B}, \mathscr{B}') = (a_{i,j})$ y $C(\mathscr{B}', \mathscr{B}'') = (b_{i,j})$ las matrices de cambio de base de $\mathscr{B}$ a $\mathscr{B}'$ y de $\mathscr{B}'$ a $\mathscr{B}''$, respectivamente. Esto significa que para cada $i \in [\![n]\!]$ es

$$x_i = a_{1,i} y_1 + \cdots + a_{n,i} y_n$$

y

$$y_i = b_{1,i} z_1 + \cdots + b_{n,i} z_n.$$

Usando esta segunda igualdad en la primera vemos que para cada $i \in [\![n]\!]$ se tiene, de hecho, que

$$\begin{aligned} x_i &= a_{1,i} y_1 + \cdots + a_{n,i} y_n \\ &= a_{1,i}(b_{1,1} z_1 + \cdots + b_{n,1} z_n) + \cdots + a_{n,i}(b_{1,n} z_1 + \cdots + b_{n,n} z_n) \\ &= (b_{1,1} a_{1,i} + \cdots + b_{1,n} a_{n,i}) z_1 + \cdots + (b_{n,1} a_{1,i} + \cdots + b_{n,n} a_{n,i}) z_n. \end{aligned}$$

Vemos así que el vector de coordenadas de x_i con respecto a la base $\mathscr{B}''$ es la i-ésima columa de la matriz $C(\mathscr{B}', \mathscr{B}'') \cdot C(\mathscr{B}, \mathscr{B}'')$ y, entonces, que esta matriz coincide con $C(\mathscr{B}, \mathscr{B}'')$, como afirma la parte (*ii*) de la proposición.

Para ver la última parte, sean $\mathscr{B}$ y $\mathscr{B}'$ dos bases ordenadas de V. De acuerdo a lo que ya probamos, se tiene que

$$C(\mathscr{B}',\mathscr{B})\cdot C(\mathscr{B},\mathscr{B}') = C(\mathscr{B},\mathscr{B}) = I_n$$

y

$$C(\mathscr{B},\mathscr{B}')\cdot C(\mathscr{B}',\mathscr{B}) = C(\mathscr{B}',\mathscr{B}') = I_n$$

y entonces es claro que las matrices $C(\mathscr{B},\mathscr{B}')$ es inversible y que que $C(\mathscr{B},\mathscr{B}')^{-1} = C(\mathscr{B}',\mathscr{B})$. La proposición que así probada. □

1.8.5. De acuerdo a la Proposición 1.8.4*(iii)*, la matriz de cambio de base entre dos bases ordenadas de un espacio vectorial de dimensión finita es una matriz inversible. Es útil muchas veces saber que, de hecho, toda matriz inversible aparece de esta forma, en el siguiente sentido:

Proposición. *Sea V un espacio vectorial de dimensión finita n. Si $\mathscr{B}$ es una base ordenada de V y $C \in \mathsf{GL}_n(\Bbbk)$ es una matriz inversible, entonces existe una y sólo una base ordenada $\mathscr{B}'$ de V tal que $C(\mathscr{B},\mathscr{B}') = C$.*

Demostración. **HACER.** □

§9. Sumas de subespacios

1.9.1. Sea V un espacio vectorial. Si $n \geq 0$ y $U_1, \dots, U_n$ son subespacios de V, entonces denotamos $U_1 + \cdots + U_n$ o $\sum_{i=1}^{n} U_i$ y llamamos ***suma de los subespacios** $U_1, \dots, U_n$* al subespacio de V generado por la unión $U_1 \cup \cdots \cup U_n$, esto es, ponemos

$$U_1 + \cdots + U_n = \langle U_1 \cup \cdots \cup U_n \rangle.$$

1.9.2. De acuerdo a la definición, la suma de una familia de subespacios es el conjunto de las combinaciones lineales de los elementos de la unión de éstos. La descripción alternativa de esta suma dada en la siguiente proposición es muchas veces útil:

Proposición. *Sea V un espacio vectorial. Si $n \geq 0$ y $U_1, \dots, U_n$ son subespacios de V, entonces*

$$U_1 + \cdots + U_n = \{x_1 + \cdots + x_n : x_1 \in U_1, \dots, x_n \in U_n\}.$$

Demostración. Escribamos T al conjunto que aparece a la derecha en la igualdad que tenemos que probar. Para ver que coincide con $U_1 + \cdots + U_n$, es decir, con $\langle U_1 \cup \cdots \cup U_n \rangle$, mostraremos que es un subespacio de V que contiene a $U_1 \cup \cdots \cup U_n$ y que es el menor subespacio de V con esta propiedad: en vista de la Proposición 1.4.6, esto es suficiente.

Primero, veamos que el subconjunto T es un subespacio de V:

- Como $0 \in T_i$ para cada $i \in [\![n]\!]$, es $0 = 0 + \cdots + 0 \in T$.

- Sean x e y dos elementos de T, de manera que para cada $i \in [\![n]\!]$ existen x_i e y_i en U_i tales que $x = x_1 + \cdots + x_n$ e $y = y_1 + \cdots + y_n$. Entonces

$$x + y = (x_1 + \cdots + x_n) + (y_1 + \cdots + y_n) = (x_1 + y_1) + \cdots + (x_n + y_n)$$

y esto nos dice que $x + y$ está en T, ya que $x_i + y_i \in U_i$ para cada $i \in [\![n]\!]$.
- Sean $a \in \Bbbk$ y $x \in T$. Existen entonces $x_1 \in U_1, \ldots, x_n \in U_n$ tales que $x = x_1 + \cdots + x_n$ y, en consecuencia,

$$a \cdot x = a \cdot (x_1 + \cdots + x_n) = ax_1 + \cdots + ax_n.$$

Como $ax_i \in U_i$ para cada $i \in [\![n]\!]$, esto muestra que $ax \in T$.

Para terminar, veamos que T es el menor subespacio de V que contiene a $U_1 \cup \cdots \cup U_n$: supongamos que W es un subespacio de V que contiene a $U_1 \cup \cdots \cup U_n$ y mostremos que necesariamente también contiene a T. Ahora bien, si $x \in T$, entonces existen $x_1 \in U_1, \ldots, x_n \in U_n$ tales que $x = x_1 + \cdots + x_n$ y, como $U_i \in W$ para cada $i \in [\![n]\!]$, esto implica que $x \in W$, ya que W es un subespacio. □

1.9.3. Proposición. *Sea V un espacio vectorial, sea $n \geq 0$ y consideremos subespacios $U_1, \ldots, U_n$ y subconjuntos $S_1, \ldots, S_n$ de V. Si cada $i \in [\![n]\!]$ el conjunto $S_i \subseteq V$ genera a U_1, entonces la unión $S_1 \cup \cdots \cup S_n$ genera a $U_1 + \cdots + U_n$.*

Demostración. Supongamos que para cada $i \in [\![n]\!]$ es $U_i = \langle S_i \rangle$ y sea $S = S_1 \cup \cdots \cup S_n$. Como para cada $i \in [\![n]\!]$ es $S_i \subseteq S$, se tiene que $U_i = \langle S_i \rangle \subseteq \langle S \rangle$. Así, es $U_1 \cup \cdots \cup U_n \subseteq \langle S \rangle$ y, por lo tanto, $U_1 + \cdots + U_n = \langle U_1 \cup \cdots \cup U_n \rangle \subseteq \langle S \rangle$. Por otro lado, como claramente $S \subseteq U_1 + \cdots + U_n$ y $U_1 + \cdots + U_n$ es un subespacio de V, es $\langle S \rangle \subseteq U_1 + \cdots + U_n$. Concluimos de esta forma que $\langle S \rangle = U_1 + \cdots + U_n$, como afirma la proposición. □

1.9.4. Proposición. *Sea V un espacio vectorial. Si U_1 y U_2 son subespacios de V de dimensión finita, entonces tanto $U_1 \cap U_2$ como $U_1 + U_2$ son subespacios de dimensión finita de V y vale que*

$$\dim(U_1 + U_2) + \dim(U_1 \cap U_2) = \dim U_1 + \dim U_2.$$

Demostración. Sean U_1 y U_2 subespacios de dimensión finita de V. Como $U_1 \cap U_2$ es un subespacio de U_1 y este último tiene dimensión finita, la Proposición 1.7.6 nos dice que $U_1 \cap U_2$ tiene dimensión finita. Sea $n = \dim U_1 \cap U_2$ y sea $\mathscr{B} = \{x_1, \ldots, x_n\}$ una base de $U_1 \cap U_2$. Como $\mathscr{B}$ es linealmente independiente y está contenido en U_1, la Proposición 1.7.6 nos dice que, si $r = \dim U_1$, es $r \geq n$ y, por otro lado, la Proposición 1.7.8 que existen vectores $y_1, \ldots, y_{r-n} \in U_1$ tales que $\mathscr{B}_1 = \{x_1, \ldots, x_n, y_1, \ldots, y_{r-n}\}$ es una base de U_1. De manera similar, si $s = \dim U_2$ es $s \geq n$ y existen vectores $z_1, \ldots, z_{s-n} \in U_2$ tales que el conjunto $\mathscr{B}_2 = \{x_1, \ldots, x_n, z_1, \ldots, z_{s-n}\}$ es una base de U_2. Mostraremos que

el conjunto $\mathscr{B}' = \{x_1, \ldots, x_n, y_1, \ldots, y_{r-n}, z_1, \ldots, z_{s-n}\}$ es una base de $U_1 + U_2$ con exactamente $r + s - n$ elementos. (14)

Por un lado, esto implicará que $U_1 + U_2$ tiene dimensión finita y, por otro, que

$$\dim(U_1 + U_2) = r + s - n = \dim U_1 + \dim U_2 - \dim(U_1 \cap U_2),$$

como afirma la proposición.

Como $\mathscr{B}' = \mathscr{B}_1 \cup \mathscr{B}_2$, la Proposición 1.9.3 nos dice que $\langle \mathscr{B}' \rangle = \langle \mathscr{B}_1 \cup \mathscr{B}_2 \rangle = U_1 + U_2$ y, en consecuencia, que el conjunto $\mathscr{B}'$ genera a $U_1 + U_2$.

Supongamos ahora que $a_1, \dots, a_n, b_1, \dots, b_{r-n}, c_1, \dots, c_{s-n} \in \Bbbk$ son escalares tales que

$$a_1x_1 + \cdots + a_nx_n + b_1y_1 + \cdots + b_{r-n}y_{r-n} + c_1z_1 + \cdots + c_{s-n}z_{s-n} = 0. \tag{15}$$

Si ponemos $x = a_1x_1 + \cdots + a_nx_n$, $y = b_1y_1 + \cdots + b_{r-n}y_{r-n}$ y $z = c_1z_1 + \cdots + c_{s-n}z_{s-n}$, la igualdad (15) nos dice que $U_1 \ni x + y = -z \in U_2$ y, entonces, que $z \in U_1 \cap U_2$. Como $\mathscr{B}$ es una base de $U_1 \cap U_2$, existen escalares $d_1, \dots, d_n \in \Bbbk$ tales que $d_1x_1 + \cdots + d_nx_n = z = c_1z_1 + \cdots + c_{s-n}z_{s-n}$, es decir,

$$d_1x_1 + \cdots + d_nx_n + (-c_1)z_1 + \cdots + (-c_{s-n})z_{s-n} = 0.$$

Ahora bien: el conjunto $\mathscr{B}_2$ es linealmente independiente, así que de esta última igualdad podemos concluir que $c_1 = \cdots = c_{s-n} = 0$. Usando esto y la ecuación (15) vemos que

$$a_1x_1 + \cdots + a_nx_n + b_1y_1 + \cdots + b_{r-n}y_{r-n} = 0$$

y, como el conjunto $\mathscr{B}_1$ es linealmente independiente, que $a_1 = \cdots = a_n = 0$ y $b_1 = \cdots = b_{r-n} = 0$. Así, todos los coeficientes que aparecen en (15) son nulos. Esto implica, primero, que los $r + s - n$ vectores $x_1, \dots, x_n, y_1, \dots, y_{r-n}, z_1, \dots, z_{s-n}$ son distintos dos a dos y, en segundo lugar, que el conjunto $\mathscr{B}'$ es linealmente independiente. Esto prueba (14), como queríamos. □

1.9.5. Corolario. *Sea V un espacio vectorial. Si U_1 y U_2 son subespacios de V de dimensión finita, entonces el subespacio $U_1 + U_2$ de V también tiene dimensión finita y*

$$\dim(U_1 + U_2) \leq \dim U_1 + \dim U_2.$$

Demostración. Esto es consecuencia inmediata de la Proposición 1.9.4, una vez que observamos que $\dim U_1 \cap U_2 \geq 0$. □

§10. Sumas directas de subespacios

1.10.1. Sea V un espacio vectorial y sean $U_1, \dots, U_k$ subespacios de V. Decimos que los subespacios $U_1, \dots, U_k$ son ***independientes*** si para cada $i \in [\![k]\!]$ se tiene que

$$U_i \cap (U_1 + \cdots + \widehat{U_i} + \cdots + U_k) = 0.$$

Si además vale que $V = U_1 + \cdots + U_k$, entonces decimos que ***V es la suma directa de $U_1, \dots, U_k$*** y escribimos $V = U_1 \oplus \cdots \oplus U_k$.

1.10.2. La siguiente observación nos permite hacer pruebas por inducción:

Lema. *Sea V un espacio vectorial y sean $U_1, \dots, U_k$ subespacios de V. Si los subespacios $U_1, \dots, U_k$ son independientes, entonces los subespacios $U_1, \dots, U_{k-1}$ son independientes y, en consecuencia, $U_1 + \cdots + U_{k-1}$ es la suma directa de $U_1, \dots, U_{k-1}$.*

Demostración. En efecto, si los subespacios $U_1, \dots, U_k$ de V son independientes e $i \in [\![k-1]\!]$, entonces

$$U_i \cap (U_1 + \cdots + \widehat{U_i} + \cdots + U_{k-1}) \subseteq U_i \cap (U_1 + \cdots + \widehat{U_i} + \cdots + U_{k-1} + U_k)$$

y esta última intersección es el subespacio nulo de V. □

1.10.3. Proposición. *Sea V un espacio vectorial y sean $U_1, \dots, U_k$ subespacios de V. Las siguientes tres afirmaciones son equivalentes:*

(a) *Es $V = U_1 \oplus \cdots \oplus U_k$.*
(b) *Para cada $x \in V$ existen únicos $x_1 \in U_1, \dots, x_k \in U_k$ tales que $x = x_1 + \cdots + x_k$.*
(c) *Es $V = U_1 + \cdots + U_k$ y si $x_1 \in U_1, \dots, x_k \in U_k$ son tales que $x_1 + \cdots + x_k = 0$, entonces $x_1 = \cdots = x_k = 0$.*

Demostración. $(a) \Rightarrow (b)$ Sea $x \in V$. Como $V = U_1 + \cdots + U_k$, la Proposición 1.9.2 nos dice que existen $x_1 \in U_1, \dots, x_k \in U_k$ tales que $x = x_1 + \cdots + x_k$. Así, vale la afirmación de existencia de (b). Para ver la unicidad, supongamos que los vectores $x_1' \in U_1, \dots, x_k' \in U_k$ también son tales que $x = x_1' + \cdots + x_l'$. Se tiene entonces que

$$(x_1 - x_1') + \cdots + (x_l - x_k') = (x_1 + \cdots + x_k) - (x_1' + \cdots + x_k') = x - x = 0.$$

Si $i \in [\![k]\!]$, esto implica que

$$-(x_i - x_i') = (x_1 - x_1') + \cdots + \widehat{(x_i - x_i')} + \cdots + (x_k - x_k')$$

y como el lado izquierdo de esta igualdad está en U_i y el lado derecho en $U_1 + \cdots + \widehat{U_i} + \cdots + U_k$, vemos que, de hecho, es

$$(x_i - x_i') \in U_i \cap (U_1 + \cdots + \widehat{U_i} + \cdots + U_k).$$

La hipótesis nos dice que esta intersección es el subespacio nulo de V, así que $x_i - x_i' = 0$, esto es, $x_i = x_i'$. Concluimos así que también vale la afirmación de unicidad de (b).

$(b) \Rightarrow (c)$ Si vale (b), entonces que $V = U_1 + \cdots + U_k$ es consecuencia inmediata de la Proposición 1.9.2 y la última afirmación de (c) es el caso particular de (b) en el que $x = 0$.

$(c) \Rightarrow (a)$ Tenemos que verificar la condición de la definición 1.10.1. Sea $i \in [\![k]\!]$ y supongamos que x es un elemento de $U_i \cap (U_1 + \cdots + \widehat{U_i} + \cdots + U_k)$. Esto significa, por un lado, que $x \in U_i$ y, por otro, que existen $x_1 \in U_1, \dots, \widehat{x_i \in U_k}, \dots, x_k \in U_k$ tales que $x = x_1 + \cdots + \widehat{x_i} + \cdots + x_k$. Como esto nos dice que

$$x_1 + \cdots + x_{i-1} + x + x_{i+1} + \cdots + x_k = 0,$$

la hipótesis (c) implica, entre otras cosas, que $x = 0$. Así, la intersección $U_i \cap (U_1 + \cdots + \widehat{U_i} + \cdots + U_k)$ sólo contiene al vector nulo y esto es lo que teníamos que probar. $\square$

1.10.4. Proposición. *Sea V un espacio vectorial. Si $U_1, \ldots, U_k$ subespacios independientes de V, todos de dimensión finita, entonces el subespacio $U_1 \oplus \cdots \oplus U_k$ de V también tiene dimensión finita y*

$$\dim(U_1 \oplus \cdots \oplus U_k) = \dim U_1 + \cdots + \dim U_k.$$

Más aún, si para cada $i \in [\![k]\!]$ el conjunto $\mathscr{B}_i$ es ua base de U_i, entonces la unión $\mathscr{B} = \mathscr{B}_1 \cup \cdots \cup \mathscr{B}_k$ es una base de $U_1 \oplus \cdots \oplus U_k$.

Demostración. Probemos la primera afirmación haciendo inducción sobre k.

- Si $k = 1$, no hay nada que probar.
- Si $k = 2$, entonces que U_1 y U_2 sean subspacios indendientes de V significa que $U_1 \cap U_2 = 0$ y entonces las dos afirmaciones del enunciado son consecuencia de la Proposición 1.9.4.
- Finalmente, sea $n \geq 2$, supongamos que la proposición es cierta si $k = n-1$ y sean $U_1, \ldots, U_n$ subespacios independientes de V. De acuerdo al Lema 1.10.2, los subespacios $U_1, \ldots, U_{n-1}$ son independientes, así que la hipótesis inductiva nos permite concluir que

  $$\dim(U_1 \oplus \cdots \oplus U_{n-1}) = \dim U_1 + \cdots + \dim U_{n-1}.$$

  Por otro lado, los subespacios $U_1 \oplus \cdots \oplus U_{n-1}$ y U_n de V son independientes, ya que la hipótesis nos dice que $(U_1 + \cdots + U_{n-1}) \cap U_n = 0$, y entonces el caso $k = 2$ de la proposición, ya que probamos, implica que

  $$\begin{aligned} \dim(U_1 \oplus \cdots \oplus U_n) &= \dim\big((U_1 \oplus \cdots \oplus U_{n-1}) \oplus U_n\big) \\ &= \dim(U_1 \oplus \cdots \oplus U_{n-1}) + \dim U_n \\ &= \dim U_1 + \cdots + \dim U_{n-1} + \dim U_n. \end{aligned}$$

  Esto completa la inducción.

Hacer: La última afirmación. $\square$

1.10.5. Proposición. *Sea V un espacio vectorial de dimensión finita y sean $U_1, \ldots, U_k$ subspacios de V. Si $V = U_1 + \cdots + U_k$ y $\dim V = \dim U_1 + \cdots + \dim U_k$, entonces $V = U_1 \oplus \cdots \oplus U_k$.*

Demostración. Hacer. $\square$

1.10.6. Proposición. *Sea V un espacio vectoria, sean $V_1, \ldots, V_r$ subespacios de V y supongamos que $V = V_1 \oplus \cdots \oplus V_r$. Si para cada $i \in [\![r]\!]$ tenemos un subespacio U_i de V_i y ponemos $U = U_1 + \cdots + U_r$, entonces, de hecho, esta suma es directa, de manera que $U = U_1 \oplus \cdots \oplus U_r$.*

Demostración. Hacer. $\square$

§11. Complementos y codimensión

1.11.1. Sea V un espacio vectorial. Si S es un subespacio de V, llamamos ***complemento de S en V*** a todo subespacio T de V tal que $V = S \oplus T$.

1.11.2. Proposición. *Sea V un espacio vectorial y sea W un subespacio de V. Si $\mathscr{B}$ es una base de V y el conjunto $\mathscr{B}' = \mathscr{B} \cap W$ genera a W, entonces $\mathscr{B}'$ es una base de W y el conjunto $\mathscr{B}'' = \mathscr{B} \smallsetminus \mathscr{B}''$ es base de un complemento de W en V.*

Demostración. **Hacer.** □

1.11.3. Corolario. *Si V es un espacio vectorial de dimensión finita, entonces todo subespacio de V posee un complemento en V.*

Demostración. **Hacer.** □

1.11.4. Proposición. *Sea V un espacio vectorial. Sea S un subespacio de V y T_1 y T_2 dos complementos de S de V. Si T_1 tiene dimensión finita, entonces T_2 también tiene dimensión finita y, de hecho,* $\dim T_1 = \dim T_2$.

Demostración. Supongamos que T tiene dimensión finita, sea $n = \dim T_1$ y sea $\mathscr{B} = \{x_1, \dots, x_n\}$ una base de T_1. Para cada $i \in [\![n]\!]$ existen $s_i \in S$ e $y_i \in T_2$ tales que $x_i = s_i + y_i$, ya que $V = S \oplus T_2$. Sea $y \in T_2$. Como $V = S \oplus T_1$, existen $s \in S$ y $x \in T_1$ tales que $y = s + x$. Como $\mathscr{B}$ es una base de T_1, existen escalares $a_1, \dots, a_n \in \Bbbk$ tales que $x = a_1x_1 + \dots + a_nx_n$. Observemos que

$$\begin{aligned} y &= s + x \\ &= s + a_1x_1 + \dots + a_nx_n \\ &= s + a_1(s_1 + y_1) + \dots + a_n(s_n + y_n) \\ &= (s + a_1s_1 + \dots + a_ns_n) + (a_1y_1 + \dots + a_ny_n), \end{aligned}$$

de manera que

$$y - (a_1y_1 + \dots + a_ny_n) = (s + a_1s_1 + \dots + a_ns_n).$$

Como el lado izquierdo de esta igualdad está en T_2 y el derecho en S, de que $S \cap T_2 = 0$ podemos deducir que, de hecho, ambos lados son nulos y, en particular, que $y = a_1y_1 + \dots + a_ny_n$. Esto nos dice que el conjunto $\{y_1, \dots, y_n\}$ genera a T_2: así, este subespacio de V tiene dimensión finita y $\dim T_2 \leq n = \dim T_1$.

Ahora que sabemos que T_2 también tiene dimensión finita podemos repetir el razonamiento que acabamos de hacer pero con los roles de T_1 y de T_2 intercambiados: concluiremos que $\dim T_1 \leq \dim T_2$ y, en definitiva, que T_1 y T_2 tienen la misma dimensión, como afirma la proposición. □

1.11.5. Si V es un espacio vectorial, decimos que un subespacio S de V tiene ***codimensión finita en V*** si posee un complemento en V de dimensión finita y en caso contrario que tiene ***codimensión infinita*** en V. Observemos que la Proposición 1.11.4 nos dice que si S tiene codimensión finita entonces todos sus complementos tienen dimensión finita y, de hecho, que todos tienen la misma dimensión: llamamos ***codimensión de S en V*** a la dimensión común de sus complementos.

1.11.6. Proposición. *Sea V un espacio vectorial. Si V tiene dimensión finita, entonces todo subespacio S de V tiene codimensión finita y*

$$\dim S + \operatorname{codim}_V S = \dim V.$$

Demostración. Supongamos que V tiene dimensión finita, sea $n = \dim V$ y sea S un subespacio de V. La Proposicion 1.7.6 que S tiene dimensión finita y que si $r = \dim S$ entonces $r \leq n$. Sea $\mathscr{B}' = \{x_1, \dots, x_r\}$ una base de S; de acuerdo a la Proposición 1.7.8, hay vectores $x_{r+1}, \dots, x_n \in V$ tales que $\mathscr{B} = \{x_1, \dots, x_n\}$ es una base de V. Consideremos el subespacio $T = \langle x_{r+1}, \dots, x_n \rangle$ de V. Como el conjunto $\mathscr{B}'' = \{x_{r+1}, \dots, x_n\}$ tiene $n-r$ elementos, estos son linealmente independientes y generan a T, es claro que $\dim T = n - r$. Para completar la prueba de la proposición bastará que mostremos que T es un complemento para S en V, esto es, que $V = S \oplus T$.

- Sea $x \in V$. Como el conjunto $\mathscr{B}$ es una base de V, hay escalares $a_1, \dots, a_n \in \Bbbk$ tales que $x = a_1x_1 + \cdots + a_nx_n$. Si ponemos $s = a_1x_1 + \cdots + a_rx_r$ y $t = a_{r+1}x_{r+1} + \cdots + a_nx_n$, entonces $s \in S$, $t \in T$ y claramente $x = s + t \in S + T$. Esto muestra que $V = S + T$.
- Por otro lado, supongamos que $x \in S \cap T$. Como $\mathscr{B}'$ y $\mathscr{B}''$ son bases de S y de T, respectivamente, existen escalares $a_1, \dots, a_r, b_1, \dots, b_{n-r} \in \Bbbk$ tales que $x = a_1x_1 + \cdots + a_rx_r$ y $x = b_1x_{r+1} + \cdots + b_{n-r}x_n$. Esto implica que

$$a_1x_1 + \cdots + a_rx_r + (-b_1)x_{r+1} + \cdots + (-b_{n-r})x_n = 0$$

  y, como $\mathscr{B}$ es linealmente independiente, que $a_1 = \cdots = a_r = 0$. Esto nos dice que $x = 0$ y, entonces, que $S \cap T = 0$.

La proposición queda así probada. □

1.11.7. Proposición. *Sea V un espacio vectorial. Si S y S' son subespacios de V tales que $S \subseteq S'$, entonces las siguientes afirmaciones son equivalentes:*

(a) *S tiene codimensión finita en V.*

(b) *S tiene codimensión finita en S' y S' tiene codimensión finita en V.*

Cuando valen, entonces

$$\operatorname{codim}_V S = \operatorname{codim}_{S'} S + \operatorname{codim}_V S'.$$

Demostración. Hacer. □

Capítulo 2

Funciones lineales

§1. Funciones lineales

2.1.1. Sean V y W espacios vectoriales. Decimos que una función $f : V \to W$ es ***lineal***, o que es un ***homomorfismo*** de espacios vectoriales, si

$$f(x+y) = f(x) + f(y), \qquad\qquad f(\lambda \cdot x) = \lambda \cdot f(x)$$

para cada $x, y \in V$ y cada $\lambda \in \mathbb{k}$. Si además es $W = V$, de manera que el dominio y codominio de f coinciden, decimos que f es un ***endomorfismo*** de V.

2.1.2. Ejemplos.

(a) Si V es un espacio vectorial, entonces la función identidad $\mathrm{id}_V : x \in V \mapsto x \in V$ es lineal. Más generalmente, si U es un subespacio de V, entonces la función de inclusión $x \in U \mapsto u \in V$ es lineal.

(b) Si V es un espacio vectorial, la función constante $x \in V \mapsto 0 \in V$ con valor el elemento cero de V es lineal.

(c) Sean m y n enteros positivos. Si $A \in \mathrm{M}_{m,n}(\mathbb{k})$ es una matriz de m filas y n columnas con entradas en $\mathbb{k}$, entonces la función $x \in \mathbb{k}^n \mapsto Ax \in \mathbb{k}^m$ es lineal. Mostraremos más adelante que, de hecho, todas las funciones lineales $\mathbb{k}^n \to \mathbb{k}^m$ son de esta forma.

(d) Sea X un conjunto y sea $\mathbb{k}^X$ el espacio vectorial de todas las funciones $X \to \mathbb{k}$. Si $x \in X$, entonces la función $f \in \mathbb{k}^X \mapsto f(x) \in \mathbb{k}$ es lineal.

(e) Sea $C(\mathbb{R})$ el espacio vectorial real de todas las funciones continuas $\mathbb{R} \to \mathbb{R}$ y sea $C^1(\mathbb{R})$ el subespacio de $C(\mathbb{R})$ de aquellas funciones que son derivables y tienen derivada continua. La función $f \in C^1(\mathbb{R}) \mapsto f' \in C(\mathbb{R})$ es lineal.

(f) Sea V un espacio vectorial de dimensión finita n y sea $\mathscr{B} = (x_1, \dots, x_n)$ una base ordenada de V. La función $x \in V \mapsto [x]_{\mathscr{B}} \in \mathbb{k}^n$ que manda cada vector de V a la n-upla ordenada de sus coordenadas con respecto a $\mathscr{B}$ es una función lineal. ◇

2.1.3. Proposición. *Sean U, V y W espacios vectoriales. Si $f : U \to V$ y $g : V \to W$ son funciones lineales, entonces la composición $g \circ f : U \to W$ también es una función lineal.*

Demostración. Sean $f : U \to V$ y $g : V \to W$ funciones lineales. Para ver que la composición $g \circ f : U \to W$ es lineal, verificamos las dos condiciones de la definición:

- Si x e y son dos vectores de U, entonces

$$\begin{aligned}(g \circ f)(x+y) &= g(f(x+y)) = g(f(x)+f(y)) = g(f(x)) + g(f(y)) \\ &= (g \circ f)(x) + (g \circ f)(y).\end{aligned}$$

- Si $\lambda \in \Bbbk$ y $x \in U$ es

$$(g \circ f)(\lambda x) = g(f(\lambda x)) = g(\lambda f(x)) = \lambda g(f(x)) = \lambda (g \circ f)(x). \qquad \square$$

2.1.4. El siguiente resultado nos da, por un lado, una forma sistemática de construir funciones lineales y, por otro, una forma de compararlas:

Proposición. *Sea V un espacio vectorial de dimensión finita n y sea $\mathscr{B} = \{x_1, \dots, x_n\}$ una base de V. Si W es un espacio vectorial e $y_1, \dots, y_n \in W$ son n elementos de W, entonces existe una y sólo una función lineal $f : V \to W$ tal que $f(x_i) = y_i$ para cada $i \in [\![n]\!]$.*

Demostración. Sea W un espacio vectorial y sean $y_1, \dots, y_n \in W$. Construimos una función $f : V \to W$ de la siguiente manera: si $x \in V$, entonces hay escalares $a_1, \dots, a_n \in \Bbbk$ unívocamente determinados por x tales que $x = a_1 x_1 + \cdots + a_n x_n$, y podemos poner

$$f(x) = a_1 y_1 + \cdots + a_n y_n.$$

Veamos que la función que obtenemos así es lineal y que satisface la condición del enunciado:

- Supongamos que x y x' son elementos de V. Sean $a_1, \dots, a_n, b_1, \dots, b_n \in \Bbbk$ escalares tales que $x = a_1 x_1 + \cdots + a_n x_n$ y $x' = b_1 x_1 + \cdots + b_n x_n$. De acuerdo a la definición de f, es entonces $f(x) = a_1 y_1 + \cdots + a_n y_n$ y $f(x') = b_1 y_1 + \cdots + b_n y_n$ y, en consecuencia,

$$\begin{aligned}f(x) + f(x') &= (a_1 y_1 + \cdots + a_n y_n) + (b_1 y_1 + \cdots + b_n y_n) \\ &= (a_1 + b_1) y_1 + \cdots + (a_n + b_n) y_n. \qquad (1)\end{aligned}$$

  Por otro lado, como

$$x + x' = (a_1 x_1 + \cdots + a_n x_n) + (b_1 x_1 + \cdots + b_n x_n) = (a_1 + b_1) x_1 + \cdots + (a_n + b_n) x_n,$$

  la definición de f nos dice que

$$f(x + x') = (a_1 + b_1) y_1 + \cdots + (a_n + b_n) y_n. \qquad (2)$$

  Como los miembros derechos de las igualdades (1) y (2) son iguales, los izquierdos también lo son y tenemos que $f(x + x') = f(x) + f(x')$.

- Sean ahora $\lambda \in \Bbbk$ y $x \in V$. Sean $a_1, \dots, a_n \in \Bbbk$ los escalares tales que $x = a_1x_1 + \cdots + a_nx_n$, de manera que $f(x) = a_1y_1 + \cdots + a_ny_n$. Como $\lambda x = \lambda a_1x_1 + \cdots + \lambda a_nx_n$, es

$$f(\lambda x) = \lambda a_1y_1 + \cdots + \lambda a_ny_n = \lambda(a_1y_1 + \cdots + a_ny_n) = \lambda f(x).$$

- Finalmente, si $i \in \llbracket n \rrbracket$, es claro que $x_i = 0x_1 + \cdots + 0x_{i-1} + 1x_i + 0x_{i+1} + \cdots + 0x_n$, así que la definición de f nos dice que $f(x_i) = 0y_1 + \cdots + 0y_{i-1} + 1y_i + 0y_{i+1} + \cdots + 0y_n = y_i$.

Con esto, la afirmación de existencia de la proposición queda probada. Veamos la unicidad.

Supongamos que $g, g' : V \to W$ son dos funciones lineales tales que $g(x_i) = y_i$ y $g'(x_i) = y_i$ para cada $i \in \llbracket n \rrbracket$. Si $x \in X$, existen escalares $a_1, \dots, a_n \in \Bbbk$ tales que $x = a_1x_1 + \cdots + a_nx_n$ y entonces, como g es lineal, se tiene que

$$g(x) = g(a_1x_1 + \cdots + a_nx_n) = a_1g(x_1) + \cdots + a_ng(x_n) = a_1y_1 + \cdots + a_ny_n.$$

De la misma forma, es $g'(x) = a_1y_1 + \cdots + a_ny_n$ y, en consecuencia, $g(x) = g'(x)$. Las funciones g y g' son, por lo tanto, iguales. $\square$

§2. Imagen y preimagen

2.2.1. Sean V y W espacios vectoriales y sea $f : V \to W$ una función lineal. Si U es un subespacio de V, la ***imagen de U por f*** es el subconjunto

$$f(U) = \{f(x) : x \in U\}$$

de W. Por otro lado, si U es un subespacio de W, la ***preimagen de U por f*** es el subconjunto

$$f^{-1}(U) = \{x \in V : f(x) \in U\}$$

de V. En particular, llamamos ***imagen de f*** y escribimos $\mathsf{Im}(f)$ a la imagen $f(V)$ de V por f, y llamamos ***núcleo de f*** y escribimos $\mathsf{Nu}(f)$ a la preimagen $f^{-1}(0)$ del subespacio nulo de W.

2.2.2. Proposición. *Sean V y W espacios vectoriales y sea $f : V \to W$ una función lineal.*

(*i*) *Si U es un subespacio de W, entonces la preimagen $f^{-1}(U)$ de U por f es un subespacio de V. En particular, el núcleo $\mathsf{Nu}(f)$ de f es un subespacio de V.*

(*ii*) *Si U es un subespacio de V, entonces la imagen $f(U)$ de U por f es un subespacio de W. En particular, la imagen $\mathsf{Im}(f)$ de f es un subespacio de W.*

Demostración. (*i*) Sea U un subespacio de W. Verifiquemos que $f^{-1}(U)$ es un subespacio de V:

- Como $f(0) = 0 \in U$, es $0 \in f^{-1}(U)$.
- Si $x, y \in f^{-1}(U)$, entonces $f(x) \in U$ y $f(y) \in U$. Como U es un subespacio de W, se sigue de esto que $f(x + y) = f(x) + f(y) \in U$, de manera que $x + y \in f^{-1}(U)$.

- Si $\lambda \in \Bbbk$ y $x \in f^{-1}(U)$, tenemos que $f(\lambda x) = \lambda f(x) \in U$, porque $f(x) \in U$, y entonces $\lambda x \in f^{-1}(U)$.

(*ii*) Sea ahora U un subespacio de V.

- Como $0 \in U$, es $0 = f(0) \in f(U)$.
- Si $x, y \in f(U)$, existen $v, w \in U$ tales que $x = f(v)$ e $y = f(w)$ y entonces

$$x + y = f(v) + f(w) = f(v + w) \in f(U),$$

  ya que $v + w \in U$.
- Si $\lambda \in \Bbbk$ y $x \in f(U)$, de manera que existe $v \in U$ tal que $x = f(v)$, es

$$\lambda x = \lambda f(v) = f(\lambda v) \in f(U)$$

  porque $\lambda v \in U$.

Esto muestra que $f(U)$ es un subespacio de W. □

§3. Monomorfismos y epimorfismos

2.3.1. Sean V y W espacios vectoriales y sea $f : V \to W$ una función lineal. Decimos que

- f es un ***monomorfismo*** si es inyectiva, y que
- f es un ***epimorfismo*** si es sobreyectiva.

2.3.2. Proposición. *Sean V y W espacios vectoriales y sea $f : V \to W$ una función lineal. Las siguientes afirmaciones son equivalentes:*

(a) *La función f es un monomorfismo.*

(b) *El núcleo $\mathsf{Nu}(f)$ de f es el subespacio nulo de V.*

(c) *Si S es un subconjunto linealmente independiente de V, entonces la imagen $f(S)$ de S por f es un subconjunto linealmente independiente de W.*

Demostración. (*a*) ⇒ (*b*) Supongamos que f es inyectiva. Si $x \in \mathsf{Nu}(f)$, entonces $f(x) = 0 = f(0)$, así que la hipótesis implica que $x = 0$: vemos que el núcleo $\mathsf{Nu}(f)$ sólo contiene al vector nulo de V y que, en consecuencia, es el subespacio nulo de V.

(*b*) ⇒ (*c*) Supongamos que $\mathsf{Nu}(f) = 0$ y probemos la afirmación contrarrecíproca de la de (*c*). Sea S un subconjunto de V y supongamos que su imagen $f(S)$ por f es linealmente dependiente. Esto significa que existen elementos $y_1, \dots, y_n \in f(S)$ distintos dos a dos y escalares $a_1, \dots, a_n \in \Bbbk$ no todos nulos tales que $a_1 y_1 + \cdots + a_n y_n = 0$. Ahora bien, como los vectores $y_1, \dots, y_n$ están en $f(S)$, existen vectores $x_1, \dots, x_n \in S$ tales que $y_i = f(x_i)$ para cada $i \in [\![n]\!]$ y, como aquéllos son distintos dos a dos, éstos también lo son. Observemos que

$$f(a_1 x_1 + \cdots + a_n x_n) = a_1 f(x_1) + \cdots + a_n f(x_n) = a_1 y_1 + \cdots + a_n y_n = 0,$$

así que $a_1x_1 + \cdots + a_nx_n \in \mathsf{Nu}(f)$. La hipótesis de que $\mathsf{Nu}(f) = 0$ nos dice entonces que, de hecho, es $a_1x_1 + \cdots + a_nx_n = 0$ y esto muestra que el conjunto S es linealmente dependiente .

$(c) \Rightarrow (a)$ Si f no es un monomorfismo, entonces existen dos vectores distintos x e y en V tales que $f(x) = f(y)$ y entonces la imagen del conjunto $S = \{x - y\}$, que es linealmente independiente, es el conjunto $f(S) = \{0\}$, que no lo es. □

2.3.3. Proposición. *Sean V y W espacios vectoriales y sea $f : V \to W$ una función lineal. Las siguientes afirmaciones son equivalentes:*

(a) *La función f es un epimorfismo.*
(b) *La imagen $\mathsf{Im}(f)$ de f es W.*
(c) *Si S es un subconjunto de V que genera a V, entonces la imagen $f(S)$ de S por f genera a W.*

Demostración. Observemos que la equivalencia entre las afirmaciones (a) y (b) es inmediata.

$(a) \Rightarrow (c)$ Supongamos que f es un epimorfismo, sea S un subconjunto de V tal que $V = \langle S \rangle$ y mostremos que $W = \langle f(S) \rangle$. Sea $y \in W$. Como f es sobreyectiva, existe $x \in V$ tal que $f(x) = y$, y como S genera a V, existen $x_1, \dots, x_n \in S$ y $a_1, \dots, a_n \in \Bbbk$ tales que $x = a_1x_1 + \cdots + a_nx_n$. Entonces

$$y = f(x) = f(a_1x_1 + \cdots + a_nx_n) = a_1f(x_1) + \cdots + a_nf(x_n) \in \langle f(S) \rangle,$$

ya que, por supuesto, los vectores $f(x_1), \dots, f(x_n)$ están en $f(S)$.

$(c) \Rightarrow (a)$ Probemos la implicación contrarrecíproca: supongamos que f no es sobreyectiva y mostremos que existe un subconjunto S de V que genera a V y tal que $f(S)$ no genera a W. De hecho, podemos poner simplemente $S = V$, que claramente genera a V. Sabemos que $f(S)$ es un subespacio de W y, como f no es sobreyectiva, que es un subespacio propio: esto nos dice que $\langle f(S) \rangle = f(S) \subsetneq W$, como queríamos. □

§4. Isomorfismos

2.4.1. Sean V y W espacios vectoriales. Una función lineal $f : V \to W$ es un ***isomorfismo*** si existe otra función lineal $g : W \to V$ tal que $g \circ f = \mathsf{id}_V$ y $f \circ g = \mathsf{id}_W$; observemos que en ese caso esta función g también es un isomorfismo, al que llamamos el ***isomorfismo inverso*** de f. Por otro lado, decimos que el espacio V es ***isomorfo*** al espacio W, y escribimos $V \cong W$, si existe un isomorfismo $V \to W$.

2.4.2. Proposición. *Sean V y W espacios vectoriales. Una función lineal $f : V \to W$ es un isomorfismo si y solamente si es biyectiva. Cuando es ése el caso, la función inversa $f^{-1} : W \to V$ es también un isomorfismo.*

Demostración. Sea $f : V \to W$ un isomorfismo y sea $g : W \to W$ el isomorfismo inverso de f, de manera que $g \circ f = \mathsf{id}_V$ y $f \circ g = \mathsf{id}_W$. Si $x \in \mathsf{Nu}(f)$, entonces $x = g(f(x)) = g(0) = 0$: esto nos dice que la función f es inyectiva. Por otro lado, si $y \in W$, entonces $y = f(g(y)) \in \mathsf{Im}(f)$:

la función f es por lo tanto sobreyectiva. Así, f es biyectiva y esto prueba la necesidad de la condición.

Veamos ahora la suficiencia. Supongamos que $f : V \to W$ es una función lineal biyectiva y mostremos que se trata de un isomorfismo. Ahora bien, como f es biyectiva, posee una función inversa $f^{-1} : W \to V$ y es suficiente que probemos que se trata de una función lineal, ya que por supuesto vale que $f^{-1} \circ f = \mathsf{id}_V$ y $f \circ f^{-1} = \mathsf{id}_W$.

- Sean x e y dos elementos de W. Como $f^{-1} \circ f = \mathsf{id}_V$ y f es lineal,

$$f(f^{-1}(x+y)) = x + y = f(f^{-1}(x) + f(f^{-1}(y)) = f(f^{-1}(x) + f^{-1}(y)),$$

  de manera que los vectores $f^{-1}(x+y)$ y $f^{-1}(x) + f^{-1}(y)$ tienen la misma imagen por f. Como f es inyectiva, esto implica que $f^{-1}(x+y) = f^{-1}(x) + f^{-1}(y)$.
- Sean $\lambda \in \Bbbk$ y $x \in W$. Usando otra vez que $f^{-1} \circ f = \mathsf{id}_V$ y que f es lineal vemos que

$$f(f^{-1}(\lambda x)) = \lambda x = \lambda f(f^{-1}(x)) = f(\lambda f^{-1}(x))$$

  y, como f es inyectiva, que $f^{-1}(\lambda x) = \lambda f^{-1}(x)$.

Esto completa la prueba de la proposición. □

2.4.3. Proposición. *La relación de isomorfismo entre espacios vectoriales es una relación de equivalencia, esto es:*

- *es* reflexiva*: cualquiera sea el espacio vectorial V, se tiene que $V \cong V$;*
- *es* simétrica*: si V y W son espacios vectoriales y $V \cong W$, entonces $W \cong V$; y*
- *es* transitiva*: si U, V y W son espacios vectoriales y se tiene que $U \cong V$ y $V \cong W$, entonces también es $U \cong W$.*

Demostración. Veamos las tres afirmaciones de a una:

- Si V es un espacio vectorial, sabemos que la función identidad $\mathsf{id}_V : x \in V \mapsto x \in V$ es lineal y, como claramente es biyectiva, se trata de un isomorfismo: se sigue de esto que $V \cong V$.
- Supongamos que V y W son espacios vectoriales y supongamos que $V \cong W$, de manera que existe un isomorfismo $f : V \to W$. De acuerdo a la Proposición 2.4.2, sabemos que la función inversa $f^{-1} : W \to V$ también es un isomorfismo y entonces $W \cong V$.
- Sean U, V y W espacios vectoriales tales que $U \cong V$ y $V \cong W$, de manera que existen isomorfismos $f : U \to V$ y $g : V \to W$. La Proposición 2.1.3 nos dice que la composición $g \circ f : U \to W$ es una función lineal. Como es además biyectiva, se trata de un isomorfismo y, por lo tanto, $U \cong V$. □

2.4.4. Proposición. *Sean V y W espacios vectoriales y sea $f : V \to W$ una función lineal. Las siguientes afirmaciones son equivalentes:*

(a) *La función f es un isomorfismo.*
(b) *Es $\mathsf{Nu}(f) = 0$ e $\mathsf{Im}(f) = W$.*
(c) *Si S es una base de V, entonces su imagen $f(S)$ por f es una base de W.*

Demostración. Esto sigue inmediatamente de las Proposiciones 2.3.2 y 2.3.3, junto con el resultado de la Proposición 2.4.2 que nos dice que una función lineal es un isomorfismo si y solamente si es a la vez un monomorfismo y un epimorfismo. □

2.4.5. Proposición. *Sean V y W espacios vectoriales.*

(i) *Si $V \cong W$ y V tiene dimensión finita, entonces W tiene dimensión finita y* $\dim V = \dim W$.

(ii) *Recíprocamente, si V y W tienen dimensión finita y* $\dim V = \dim W$, *entonces $V \cong W$.*

Demostración. (i) Si $V \cong W$ y V tiene dimensión finita, entonces hay un isomorfismo $f : V \to W$ y una base finita $\mathscr{B}$ de V. De acuerdo a la Proposición 2.4.4, el conjunto $f(\mathscr{B})$ es una base de W y, como es finita, W tiene dimensión finita. Más precisamente, como f es una biyección, $\mathscr{B}$ y $f(\mathscr{B})$ tienen la misma cantidad de elementos, así que $\dim V = \dim W$.

(ii) Supongamos que V y W tienen dimensión finita y que $\dim V = \dim W$, y sea $n = \dim V$. Sean $\mathscr{B} = \{x_1, \dots, x_n\}$ y $\mathscr{B}' = \{y_1, \dots, y_n\}$ bases de V y de W, respectivamente. La Proposición 2.1.4 nos dice que existen funciones lineales $f : V \to W$ y $g : W \to V$ tales que $f(x_i) = y_1$ y $g(y_i) = x_i$ para cada $i \in [\![n]\!]$.

La función $g \circ f : V \to V$ es lineal y es tal que $(g \circ f)(x_i) = x_i$ para cada $i \in [\![n]\!]$: se trata de una función lineal $V \to V$ que coincide sobre la base $\mathscr{B}$ de V con la función identidad $\mathrm{id}_V : V \to V$, y la Proposición 2.1.4 implica entonces que $g \circ f = \mathrm{id}_V$. De la misma forma, podemos ver que $f \circ g = \mathrm{id}_W$ y, en consecuencia, concluir que f es un isomorfismo. Así, V y W son espacios vectoriales isomorfos, como afirma la proposición. □

§5. El teorema de la dimensión

2.5.1. Teorema. *Sean V y W dos espacios vectoriales y sea $f : V \to W$ una función lineal. Son equivalentes las siguientes afirmaciones:*

(a) *El espacio vectorial V tiene dimensión finita.*

(b) *Tanto el núcleo* $\mathrm{Nu}(f)$ *como la imagen* $\mathrm{Im}(f)$ *de f tienen dimensión finita.*

Cuando estas afirmaciones se cumplen, se tiene además que

$$\dim V = \dim \mathrm{Nu}(f) + \dim \mathrm{Im}(f). \tag{3}$$

Demostración. Supongamos primero que V tiene dimensión finita y sea $n = \dim V$. Como $\mathrm{Nu}(f)$ es un subespacio de V, la Proposición 1.7.6 nos dice que tiene dimensión finita. Sea r su dimensión y sea $\{x_1, \dots, x_r\}$ una de sus bases. Como $\mathrm{Nu}(f)$ está contenido en V, este conjunto $\{x_1, \dots, x_r\}$ está contenido en V y es allí linealmente independiente. La Proposición 1.7.8 nos dice que existen vectores $x_{r+1}, \dots, x_n \in V$ tales que el conjunto $\mathscr{B} = \{x_1, \dots, x_r, x_{r+1}, \dots, x_n\}$ es una base de V. Mostraremos que

el subconjunto $\mathscr{B}' = \{f(x_{r+1}), \dots, f(x_n)\}$ de W tiene $n - r$ elementos y es una base de $\mathrm{Im}(f)$. (4)

Se sigue de esto que $\mathsf{Im}(f)$ tiene dimensión finita, que $\dim \mathsf{Im}(f) = n - r$ y, en consecuencia, que

$$\dim V = n = r + (n-r) = \dim \mathsf{Nu}(f) + \dim \mathsf{Im}(f).$$

Esto probará que la afirmación (*a*) del teorema implica la afirmación (*b*) y que cuando vale aquella vale la igualdad (3).

Veamos entonces (4) Sea primero $y \in \mathsf{Im}(f)$, de manera que existe $x \in V$ tal que $y = f(x)$. Como $\mathscr{B}$ es una base de V, existen escalares $a_1, \dots, a_n \in \Bbbk$ tales que $x = a_1x_1 + \cdots + a_nx_n$ y entonces

$$\begin{aligned} y = f(x) &= f(a_1x_1 + \cdots + a_nx_n) \\ &= b_1f(x_1) + \cdots + a_nf(x_n) \\ &= a_{r+1}f(x_{r+1}) + \cdots + a_nf(x_n) \in \langle \mathscr{B}' \rangle, \end{aligned}$$

porque $f(x_1) = \cdots = f(x_r) = 0$. Vemos así que el conjunto $\mathscr{B}'$ genera a $\mathsf{Im}(f)$.

Por otro lado, supongamos que $b_1, \dots, b_{n-r} \in \Bbbk$ son escalares tales que

$$b_1f(x_{r+1}) + \cdots + b_{n-r}f(x_n) = 0.$$

Como f es lineal, esto nos dice que $f(b_1x_{r+1} + \cdots + b_{n-r}x_n) = 0$ y, en consecuencia, que el vector $b_1x_{r+1} + \cdots + b_{n-r}x_n$ está en el núcleo $\mathsf{Nu}(f)$. Como el conjunto $\{x_1, \dots, x_r\}$ es una base de este subespacio, existen entonces escalares $c_1, \dots, c_r \in \Bbbk$ tales que

$$c_1x_1 + \cdots + c_rx_r = b_1x_{r+1} + \cdots + b_{n-r}x_n.$$

Ahora bien: como el conjunto $\mathscr{B}$ es una base, es linealmente independiente y esta última igualdad implica que, entre otras cosas, $b_1 = \cdots = b_{n-r} = 0$. Concluimos de esta forma que los $n - r$ vectores $f(x_{r+1}), \dots, f(x_n)$ son distintos dos a dos y que el conjunto $\mathscr{B}'$ es linealmente independiente.

Para completar la prueba del teorema, tenemos que mostrar que vale la implicación (*b*) $\Rightarrow$ (*a*). Supongamos que $\mathsf{Nu}(f)$ e $\mathsf{Im}(f)$ tienen dimensión finita, sean $r = \dim \mathsf{Nu}(f)$ y $s = \dim \mathsf{Im}(f)$ y sean $\mathscr{B}' = \{x_1, \dots, x_r\}$ y $\mathscr{B}'' = \{y_1, \dots, y_s\}$ bases de $\mathsf{Nu}(f)$ y de $\mathsf{Im}(f)$, respectivamente. Si $i \in [\![s]\!]$, el vector y_i está en $\mathsf{Im}(f)$, así que existe $z_i \in V$ tal que $f(z_i) = y_i$. En vista de la Proposición 1.6.3, para ver que V tiene dimensión finita bastará que veamos que el conjunto $\mathscr{B} = \{x_1, \dots, x_r, z_1, \dots, z_s\}$ genera a V.

Sea entonces $x \in V$. Como $f(x) \in \mathsf{Im}(f)$ y el conjunto $\mathscr{B}''$ es una base de $\mathsf{Im}(f)$, existen escalares $b_1, \dots, b_s \in \Bbbk$ tales que $f(x) = b_1y_1 + \cdots + b_sy_s$ y entonces

$$\begin{aligned} f\big(x - (b_1z_1 + \cdots + b_sz_s)\big) &= f(x) - f(b_1z_1 + \cdots + b_sz_s) \\ &= f(x) - \big(b_1f(z_1) + \cdots + b_sf(z_s)\big) \\ &= f(x) - (b_1y_1 + \cdots + b_sy_s) \\ &= f(x) - f(x) \\ &= 0. \end{aligned}$$

Así, el vector $x - (b_1z_1 + \cdots + b_sz_s)$ está en $\mathsf{Nu}(f)$ y, como $\mathscr{B}'$ es una base de ese subespacio, existen escalares $a_1, \ldots, a_r \in \Bbbk$ tales que

$$x - (b_1z_1 + \cdots + b_sz_s) = a_1x_1 + \cdots + a_rx_r.$$

Por supuesto, esto nos dice que

$$x = a_1x_1 + \cdots + a_rx_r + b_1z_1 + \cdots + b_sz_s \in \langle \mathscr{B} \rangle$$

y prueba, en definitiva, que $\mathscr{B}$ genera a V, como queríamos. □

2.5.2. Proposición. *Sean V y W espacios vectoriales y sea $f : V \to W$ una función lineal.*

(*i*) *Si f es sobreyectiva y V tiene dimensión finita, entonces W también tiene dimensión finita y* $\dim W \leq \dim V$.

(*ii*) *Si f es inyectiva y W tiene dimensión finita, entonces V también tiene dimensión finita y* $\dim V \leq \dim W$.

Demostración. (*i*) Supongamos que f es sobreyectiva y que V tiene dimensión finita. Esto último, de acuerdo al Teorema 2.5.1, implica que $\mathsf{Nu}(f)$ y $W = \mathsf{Im}(f)$ tienen dimensión finita y que

$$\dim V = \dim \mathsf{Nu}(f) + \dim \mathsf{Im}(f) \geq \dim \mathsf{Im}(f) = \dim W.$$

(*ii*) Supongamos ahora que f es inyectiva y que W tiene dimensión finita. La imagen $\mathsf{Im}(f)$ de f, que es un subespacio de W, tiene entonces dimensión finita. Por otro lado, como la función f es inyectiva, su núcleo $\mathsf{Nu}(f)$ es el subespacio nulo de V, que tiene dimensión finita —nula, de hecho— y el Teorema 2.5.1 nos dice que V tiene dimensión finita y que

$$\dim V = \dim \mathsf{Nu}(f) + \dim \mathsf{Im}(f) = \dim \mathsf{Im}(f) \leq \dim W,$$

como afirma el enunciado. □

2.5.3. Proposición. *Sean V y W espacios vectoriales y sea $f : V \to W$ una función lineal. Si V y W tienen dimensión finita y* $\dim V = \dim W$, *entonces las siguientes afirmaciones son equivalentes:*

(a) *f es un isomorfismo.*

(b) *f es un monomorfismo.*

(c) *f es un epimorfismo.*

Demostración. Supongamos que los espacios V y W tienen dimensión finita y que $\dim V = \dim W$. Es claro que si f es un isomorfismo, entonces es un monomorfismo y un epimorfismo: esto nos dice que (*a*) implica a (*b*) y a (*c*). Veamos las implicaciones recíprocas. Del Teorema 2.5.1 sabemos que

$$\dim V = \dim \mathsf{Nu}(f) + \dim \mathsf{Im}(f). \tag{5}$$

Si f es un monomorfismo, entonces el núcleo de f es el subespacio nulo de V y la igualdad (5) nos dice que $\dim W = \dim V = \dim \mathsf{Im}(f)$: como $\mathsf{Im}(f)$ es un subespacio de W, esto implica que $\mathsf{Im}(f) = W$, esto es, que f es sobreyectiva. Si, en cambio, f es un epimorfismo, entonces $\mathsf{Im}(f) = W$, de manera que $\dim \mathsf{Im}(f) = \dim W = \dim V$ y, de acuerdo a (5), $\dim \mathsf{Nu}(f) = 0$. Esto significa que $\mathsf{Nu}(f) = 0$ y, en consecuencia, que f es inyectiva. $\square$

§6. El espacio de homomorfismos entre dos espacios vectoriales

2.6.1. Si W son dos espacios vectoriales, escribimos $\hom(V, W)$ al conjunto de las funciones lineales $V \to W$. Es un subconjunto del espacio vectorial W^V de *todas* las funciones $V \to W$ que construimos en el Ejemplo 1.2.4(b) y, de hecho, se trata de un subespacio:

- El elemento nulo de W^V es la función constante nula $x \in V \mapsto 0 \in W$ y esta función es lineal, por lo que pertenece a $\hom(V, W)$.
- Si $f, g \in \hom(V, W)$, entonces $f + g \in \hom(V, W)$: en efecto, para cada $x, y \in V$ tenemos que

$$\begin{aligned}(f+g)(x+y) &= f(x+y) + g(x+y) \\ &= f(x) + f(y) + g(x) + g(y) \\ &= f(x) + g(x) + f(y) + g(y) \\ &= (f+g)(x) + (f+g)(y)\end{aligned}$$

  y para cada $\lambda \in \Bbbk$ y $x \in V$, que

$$\begin{aligned}(f+g)(\lambda x) &= f(\lambda x) + g(\lambda x) \\ &= \lambda f(x) + \lambda g(x) \\ &= \lambda\big(f(x) + g(x)\big) \\ &= \lambda(f+g)(x).\end{aligned}$$

- Si $\lambda \in \Bbbk$ y $f \in \hom(V, W)$ entonces $\lambda f \in \hom(V, W)$, ya que para cada $x, y \in V$ es

$$\begin{aligned}(\lambda f)(x+y) &= \lambda f(x+y) \\ &= \lambda\big(f(x) + f(y)\big) \\ &= \lambda f(x) + \lambda f(y) \\ &= (\lambda f)(x) + (\lambda f)(y)\end{aligned}$$

  y para cada $\mu \in \Bbbk$ y $x \in V$, que

$$\begin{aligned}(\lambda f)(\mu x) &= \lambda f(\mu x) \\ &= \lambda\mu f(x) \\ &= \mu\lambda f(x) \\ &= \mu(\lambda f)(x).\end{aligned}$$

Desde ahora en adelante, cada vez que veamos a $\mathsf{hom}(V, W)$ lo consideraremos con la estructura de espacio vectorial que tiene como subespacio de W^V.

2.6.2. Si U, V y W son espacios vectoriales, decimos que una función $\beta : V \times W \to U$ es ***bilineal*** si

$$\beta(x + x', y) = \beta(x, y) + \beta(x', y), \qquad \beta(\lambda x, y) = \lambda\beta(x, y),$$
$$\beta(x, y + y') = \beta(x, y) + \beta(x, y'), \qquad \beta(x, \lambda y) = \lambda\beta(x, y)$$

para cada x, $x' \in V$, y, $y' \in W$, $\lambda \in \Bbbk$. Observemos que las primeras dos condiciones nos dicen que para todo $y \in W$ la función $x \in V \mapsto \beta(x, y) \in U$ es lineal, mientras que las dos últimas condiciones que para todo $x \in V$ la función $y \in W \mapsto \beta(x, y) \in U$ es lineal: decimos por eso que una función bilineal es una que es lineal en cada una de sus dos variables.

2.6.3. Proposición. *Sean V y W dos espacios vectoriales. La función*

$$\varepsilon : (f, x) \in \mathsf{hom}(V, W) \times V \mapsto f(x) \in W$$

es bilineal.

Demostración. Sean f, $f' \in \mathsf{hom}(V, W)$, x, $x' \in V$ y $\lambda \in \Bbbk$. Calculando, vemos que

$$\varepsilon(f + f', x) = (f + f')(x) = f(x) + f'(x) = \varepsilon(f, x) + \varepsilon(f', x),$$
$$\varepsilon(\lambda f, x) = (\lambda f)(x) = \lambda f(x) = \lambda\varepsilon(f, x),$$
$$\varepsilon(f, x + x') = f(x + x') = f(x) + f(x') = \varepsilon(f, x) + \varepsilon(f, x')$$

y

$$\varepsilon(f, \lambda x) = f(\lambda x) = \lambda f(x) = \lambda\varepsilon(f, x),$$

y estas cuatro igualdades nos dicen que la función ε es bilineal. □

2.6.4. Proposición. *Sean U, V y W espacios vectoriales. La función*

$$\beta : (f, g) \in \mathsf{hom}(V, W) \times \mathsf{hom}(U, V) \mapsto f \circ g \in \mathsf{hom}(U, W)$$

es bilineal.

Demostración. Sean f, $f' \in \mathsf{hom}(V, W)$, $g \in \mathsf{hom}(U, V)$ y $\lambda \in \Bbbk$, y verifiquemos las dos primeras condiciones de la definición:

- Para cada $x \in U$ es

$$\begin{aligned}
\beta(f + f', g)(x) &= \big((f + f') \circ g\big)(x) \\
&= (f + f')\big(g(x)\big) \\
&= f(g(x)) + f'(g(x)) \\
&= (f \circ g)(x) + (f' \circ g)(x) \\
&= \beta(f, g)(x) + \beta(f', g)(x) \\
&= \big(\beta(f, g) + \beta(f', g)\big)(x)
\end{aligned}$$

así que $\beta(f+f',g)=\beta(f,g)+\beta(f',g)$,

- Para cada $x\in U$, calculamos que

$$\begin{aligned}\beta(\lambda f,g)(x)&=\big((\lambda f)\circ g\big)(x)\\&=(\lambda f)\big(g(x)\big)\\&=\lambda f(g(x))\\&=\lambda(f\circ g)(x)\\&=\lambda\beta(f,g)(x),\end{aligned}$$

así que $\beta(\lambda f,g)=\lambda\beta(f,g)$.

La verificación de las otras dos condiciones es enteramente similar. □

§7. El álgebra de endomorfismos de un espacio vectorial

2.7.1. Un ***álgebra sobre*** $\Bbbk$ es un espacio vectorial A dotado de una operación de multiplicación

$$\cdot : A\times A\to A$$

tal que

($\mathbf{A}_1$) la multiplicación $\cdot$ es una función bilineal;

($\mathbf{A}_2$) la multiplicación $\cdot$ es asociativa: para cada a, b, $c\in A$ se tiene que $(a\cdot b)\cdot c=a\cdot(b\cdot c)$;

($\mathbf{A}_3$) la multiplicación $\cdot$ posee un elemento neutro: existe un elemento $1\in A$ tal que para cada $a\in A$ se tiene que $a\cdot 1=a$ y $1\cdot a=a$.

Es importante observar que no exigimos que la multiplicación de un álgebra sea conmutativa y, de hecho, en general no lo es.

2.7.2. Ejemplos.

(a) Veamos al cuerpo $\Bbbk$ con su estructura usual de espacio vectorial sobre $\Bbbk$. Si dotamos a $\Bbbk$ de la multiplicación $\cdot:\Bbbk\times\Bbbk\to\Bbbk$ que tiene en tanto cuerpo, $\Bbbk$ es un álgebra sobre $\Bbbk$.

(b) Más generalmente, si K es un cuerpo y $\Bbbk\subseteq K$ es un subcuerpo de K, vimos en el Ejemplo 1.2.5(c) que podemos ver a K como un espacio vectorial sobre $\Bbbk$ de manera natural y, de hecho, si lo dotamos de su multiplicación, resulta ser un álgebra sobre $\Bbbk$.

(c) Sea $\Bbbk[X]$ el espacio vectorial de los polinomios en la variable X con coeficientes en $\Bbbk$. La multiplicación $\cdot:\Bbbk[X]\times\Bbbk[X]\to\Bbbk[X]$ es bilineal, asociativa y posee un elemento neutro, así que hace de $\Bbbk[X]$ un álgebra sobre $\Bbbk$.

(d) Sea X un conjunto no vacío y sea $\Bbbk^X$ el espacio vectorial de todas las funciones $X\to\Bbbk$, como en el Ejemplo 1.2.4(a). Si dotamos a $\Bbbk^X$ de la operación usual $\cdot:\Bbbk^X\times\Bbbk^X\to\Bbbk^X$ de multiplicación de funciones, entonces es un álgebra sobre $\Bbbk$. ◇

2.7.3. Los dos ejemplos de álgebras más importantes para nosotros son los que da la siguiente proposición.

Proposición.

(i) *Sea V un espacio vectorial. El espacio vectorial* $\mathsf{End}(V)$ *de los endomorfismos de V es un álgebra sobre $\Bbbk$ con respecto a la operación de multiplicación*

$$\cdot : (f, g) \in \mathsf{End}(V) \times \mathsf{End}(V) \mapsto f \circ g \in \mathsf{End}(V)$$

dada por la composición.

(ii) *Sea $n \in \mathbb{N}$. El espacio vectorial* $\mathsf{M}_n(\Bbbk)$ *de las matrices cuadradas de n filas y n columnas es un álgebra sobre $\Bbbk$ con respecto a la operación de multiplicación matricial*

$$\cdot : (A, B) \in \mathsf{M}_n(\Bbbk) \times \mathsf{M}_n(\Bbbk) \mapsto AB \in \mathsf{M}_n(\Bbbk).$$

Demostración. (i) Sabemos de la Proposición 2.6.4 que esta operación es bilineal. Es asociativa porque la composición de funciones es una operación asociativa y claramente tiene a la función identidad $\mathsf{id}_V \in \mathsf{End}(V)$ como elemento neutro.

(ii) Esto es consecuencia de la distributividad y la asociatividad de la multiplicación matricial, y de que la matriz identidad I_n es un elemento neutro para la multiplicación. □

§8. La matriz asociada a una función lineal

2.8.1. Sean V y W dos espacios vectoriales de dimensión finita, sean $n = \dim V$ y $m = \dim W$ sus dimensiones, y sean $\mathscr{B} = (x_1, \dots, x_n)$ y $\mathscr{B}' = (y_1, \dots, y_m)$ bases ordenadas de V y de W, respectivamente. Si $f : V \to W$ es una función lineal, para cada $i \in [\![n]\!]$ hay escalares $a_{1,i}, \dots, a_{m,i} \in \Bbbk$ bien determinados tales que

$$f(x_i) = a_{1,i} y_i + \cdots + a_{m,i} y_m,$$

las coordenadas de $f(x_i)$ con respecto a la base $\mathscr{B}'$. De esta manera obtenemos mn elementos de $\Bbbk$, que podemos ordenar en forma de una matriz

$$\begin{pmatrix} a_{1,1} & \cdots & a_{1,n} \\ \vdots & \ddots & \vdots \\ a_{m,1} & \cdots & a_{m,n} \end{pmatrix} \in \mathsf{M}_{m,n}(\Bbbk)$$

de m filas y n columnas con entradas en $\Bbbk$, a la que llamamos ***la matriz de f con respecto a las bases ordenadas $\mathscr{B}$ y $\mathscr{B}'$*** y escribimos $[f]^{\mathscr{B}}_{\mathscr{B}'}$. Explícitamente: para cada $i \in [\![n]\!]$, la i-ésima columna de esta matriz tiene las coordenadas, en orden, de la imagen del i-ésimo vector de la base $\mathscr{B}$ con respecto a la base $\mathscr{B}'$.

2.8.2. Ejemplos.

(a) Sean m, $n \in \mathbb{N}$, sea $A = (a_{i,j}) \in \mathsf{M}_{m,n}(\Bbbk)$ una matriz y consideremos la función lineal $f_A : x \in \Bbbk^n \mapsto Ax \in \Bbbk^m$. Sean $\mathscr{B} = (x_1, \dots, x_n)$ y $\mathscr{B}' = (y_1, \dots, y_m)$ son las bases ordenadas estándares de $\Bbbk^n$ y de $\Bbbk^m$, respectivamente. Como el vector $f_A(x_i) = Ax_i$ es la columna i-ésima de la matriz A, esto es, $(a_{1,i}, \dots, a_{m,i})^{\mathsf{t}}$, y este vector se escribe en la forma

$$f_A(x_i) = a_{1,i}y_1 + \cdots + a_{m,i}y_m$$

en la base $\mathscr{B}'$, vemos que la matriz de f_A con respecto a las bases ordenadas $\mathscr{B}$ y $\mathscr{B}'$ es precisamente la matriz A.

(b) Sea V un espacio vectorial de dimensión finita y sea $n = \dim V$. Si $\mathrm{id}_V : V \to V$ es la función identidad de V, entonces para cada base ordenada $\mathscr{B}$ de V se tiene que

$$[\mathrm{id}_V]^{\mathscr{B}}_{\mathscr{B}} = I_n = \begin{pmatrix} 1 & 0 & \cdots & 0 \\ 0 & 1 & & \vdots \\ \vdots & & \ddots & 0 \\ 0 & \cdots & 0 & 1 \end{pmatrix},$$

la matriz identidad de n filas y n columnas. Observemos que es es cierto porque estamos usando la misma base $\mathscr{B}$ tanto para el dominio como para el codominio de id_V: si usásemos bases distintas, la matriz que obtendríamos sería otra.

(c) Sea $n \in \mathbb{N}$ y consideremos el espacio vectorial $V = \Bbbk[X]_{\leq n}$ de los polinomios con coeficientes en $\Bbbk$ de grado a lo sumo n. Sea $f : p \in V \mapsto p' \in V$ el endomorfismo de V que manda cada elemento de V a su derivada y sea $\mathscr{B} = \{1, X, \dots, X^n\}$ la base usual de V. Sabemos que $f(1) = 0$ y que para cada $i \in [\![n]\!]$ es $f(X^i) = iX^{i-1}$. Esto implica inmediatamente que la matriz de f con respecto a las bases $\mathscr{B}$ de su dominio y $\mathscr{B}$ de su codominio es la matriz cuadrada de $n+1$ filas y columnas

$$[f]^{\mathscr{B}}_{\mathscr{B}} = \begin{pmatrix} 0 & 1 & 0 & & \cdots & 0 \\ & 0 & 2 & 0 & & \vdots \\ & & \ddots & \ddots & \ddots & \\ & & & 0 & n-1 & 0 \\ \vdots & & & & 0 & n \\ 0 & \cdots & & & & 0 \end{pmatrix}. \qquad \diamond$$

2.8.3. La razón por la que estamos interesados en la matriz de una función lineal es la siguiente:

Proposición. *Sean V y W dos espacios vectoriales de dimensión finita y sean $\mathscr{B}$ y $\mathscr{B}'$ bases ordenadas de V y de W, respectivamente. Si $f : V \to W$ es una función lineal y $x \in V$, entonces*

$$[f(x)]_{\mathscr{B}'} = [f]^{\mathscr{B}}_{\mathscr{B}'} \cdot [x]_{\mathscr{B}}$$

y, de hecho, $[f]^{\mathscr{B}}_{\mathscr{B}'}$ es la única matriz con esta propiedad.

En otras palabras: el vector de coordenadas del vector $f(x)$ con respecto a la base ordenada $\mathscr{B}'$ es el producto de la matriz $[f]_{\mathscr{B}}^{\mathscr{B}'}$ de f y el vector de coordenadas $[x]_{\mathscr{B}}$ de x con respecto a la base $\mathscr{B}$.

Demostración. Sean $n = \dim V$ y $m = \dim W$, y sean $\mathscr{B} = (x_1, \dots, x_n)$ y $\mathscr{B}' = (y_1, \dots, y_m)$. Sea $f : V \to W$ una función lineal y sea $x \in V$. Si $[x]_{\mathscr{B}} = (a_1, \dots, a_n)^{\mathsf{t}} \in \Bbbk^n$ es el vector de coordenadas de x y $[f]_{\mathscr{B}'}^{\mathscr{B}} = (b_{i,j}) \in \mathsf{M}_{m,n}(\Bbbk)$ es la matriz de f, entonces

$$x = a_1 x_1 + \cdots + a_n x_n$$

y

$$f(x_j) = b_{1,j} y_1 + \cdots + b_{m,j} y_m$$

para cada $j \in [\![n]\!]$ y, en consecuencia,

$$\begin{aligned} f(x) &= f(a_1 x_1 + \cdots + a_n x_n) \\ &= a_1 f(x_1) + \cdots + a_n f(x_n) \\ &= a_1(b_{1,1} y_1 + \cdots + b_{m,1} y_m) + \cdots + a_n(b_{1,n} y_1 + \cdots + b_{m,n} y_m) \\ &= (b_{1,1} a_1 + \cdots + b_{1,n} a_n) y_1 + \cdots + (b_{m,1} a_1 + \cdots + b_{m,n} a_n) y_n. \end{aligned}$$

Esto significa que

$$[f(x)]_{\mathscr{B}'} = \begin{pmatrix} b_{1,1} a_1 + \cdots + b_{1,n} a_n \\ \vdots \\ b_{m,1} a_1 + \cdots + b_{m,n} a_n \end{pmatrix} = \begin{pmatrix} b_{1,1} & \cdots & b_{1,n} \\ \vdots & \ddots & \vdots \\ b_{m,1} & \cdots & b_{m,n} \end{pmatrix} \begin{pmatrix} a_1 \\ \cdots \\ a_n \end{pmatrix} = [f]_{\mathscr{B}'}^{\mathscr{B}} \cdot [x]_{\mathscr{B}},$$

como afirma la proposición.

Para ver la unicidad, supongamos que $A \in \mathsf{M}_{m,n}(\Bbbk)$ es otra matriz tal que $[f(x)]_{\mathscr{B}'} = A \cdot [x]_{\mathscr{B}}$ para todo $x \in V$. Si $i \in [\![n]\!]$, entonces la columna i-ésima de A es, escribiendo e_i al vector i-ésimo de la base ordenada estándar de $\Bbbk^n$,

$$A \cdot e_i = A \cdot [x_i]_{\mathscr{B}} = [f(x_i)]_{\mathscr{B}'},$$

y este último vector es precisamente la columa i-ésima de $[f]_{\mathscr{B}'}^{\mathscr{B}}$. Esto muestra que, de hecho, $A = [f]_{\mathscr{B}'}^{\mathscr{B}}$. □

2.8.4. Proposición. *Sean V y W dos espacios vectoriales de dimensión finita, sean $n = \dim V$ y $m = \dim W$ sus dimensiones, y sean $\mathscr{B}$ y $\mathscr{B}'$ bases ordenadas de V y de W, respectivamente. La función*

$$\Phi : f \in \hom(V, W) \mapsto [f]_{\mathscr{B}'}^{\mathscr{B}} \in \mathsf{M}_{m,n}(\Bbbk)$$

es un isomorfismo de espacios vectoriales. En particular, el espacio vectorial $\hom(V, W)$ *tiene dimensión finita y*

$$\dim \hom(V, W) = \dim V \cdot \dim W.$$

Demostración. Sean $\mathscr{B} = (x_1, \dots, x_n)$ y $\mathscr{B}' = (y_1, \dots, y_m)$ bases ordenadas de V y de W, respectivamente. Mostremos primero que la función Φ es lineal:

- Sean $f, g \in \mathsf{hom}(V, W)$, y supongamos que $\Phi(f) = [f]^{\mathscr{B}}_{\mathscr{B}'} = (a_{i,j})$ y $\Phi(f) = [g]^{\mathscr{B}}_{\mathscr{B}'} = (b_{i,j})$. Esto significa que para cada $i \in [\![n]\!]$ es

  $$f(x_i) = a_{1,i}y_1 + \cdots + a_{m,i}y_m$$

  y

  $$g(x_i) = b_{1,i}y_1 + \cdots + b_{m,i}y_m,$$

  y entonces

  $$\begin{aligned}(f+g)(x_i) &= f(x_i) + g(x_i)\\ &= (a_{1,i}y_1 + \cdots + a_{m,i}y_m) + (b_{1,i}y_1 + \cdots + b_{m,i}y_m)\\ &= (a_{1,i} + b_{1,i})y_1 + \cdots + (a_{m,i} + b_{m,i})y_m.\end{aligned}$$

  Vemos así que

  $$\Phi(f+g) = [f+g]^{\mathscr{B}}_{\mathscr{B}'} = (a_{i,j} + b_{i,j}) = \Phi(f) + \Phi(g).$$

- Sean ahora $\lambda \in \Bbbk$ y $f \in \mathsf{hom}(V, W)$. Sea $\Phi(f) = [f]^{\mathscr{B}}_{\mathscr{B}'} = (a_{i,j})$, de manera que para cada $i \in [\![n]\!]$ es $f(x_i) = a_{1,i}y_1 + \cdots + a_{m,i}y_m$. Como entonces

  $$\begin{aligned}(\lambda f)(x_i) &= \lambda f(x_i)\\ &= \lambda(a_{1,i}y_1 + \cdots + a_{m,i}y_m)\\ &= (\lambda a_{1,i})y_1 + \cdots + (\lambda a_{m,i})y_m,\end{aligned}$$

  se tiene que

  $$\Phi(\lambda f) = [\lambda f]^{\mathscr{B}}_{\mathscr{B}'} = (\lambda a_{i,j}) = \lambda(a_{i,j}) = \lambda\Phi(f).$$

Para ver que la función Φ es inyectiva, mostremos que tiene núcleo nulo. Sea $f \in \mathsf{hom}(V, W)$ y supongamos que $\Phi(f)$ es el elemento nulo de $\mathsf{M}_{m,n}(\Bbbk)$, esto es, que la matriz $[f]^{\mathscr{B}}_{\mathscr{B}'} = (a_{i,j})$ tiene todas sus entradas nulas. Esto significa que $f(x_i) = 0$ para cada $i \in [\![n]\!]$ y, entonces, que f coincide con la función nula $0 : V \to W$ en cada elemento de la base $\mathscr{B}$. Sabemos, de la Proposición 2.1.4 que esto implica que $f = 0$. Así, $\mathsf{Nu}(\Phi) = 0$, como queríamos.

Por otro lado, la función Φ es sobreyectiva: si $A = (a_{i,j})$ es un elemento de $\mathsf{M}_{m,n}(\Bbbk)$, la Proposición 2.1.4 nos dice que existe una función lineal $f : V \to W$ tal que $f(x_i) = a_{1,i}y_1 + \cdots + a_{m,i}y_m$ para cada $i \in [\![n]\!]$ y es claro que esta función tiene matriz $[f]^{\mathscr{B}}_{\mathscr{B}'} = A$. Así, $\Phi(f) = A$ y A está en la imagen de Φ.

Hemos probado así que Φ es un isomorfismo de espacios vectoriales. Como el espacio de matrices $\mathsf{M}_{m,n}(\Bbbk)$ tiene dimensión finita y, de hecho, su dimensión en mn, la Proposición 2.4.5 nos dice que $\mathsf{hom}(V, W)$ tiene dimensión finita y que su dimensión es mn. Esto completa la prueba de la proposición. □

2.8.5. Proposición. *Sean U, V y W tres espacios vectoriales de dimensión finita y sean $\mathscr{B}$, $\mathscr{B}'$ y $\mathscr{B}''$ bases ordenadas de U, de V y de W, respectivamente. Si $f : U \to V$ y $g : V \to W$ son funciones lineales, entonces la matriz de la composición $g \circ f : U \to W$ con respecto a las bases ordenadas $\mathscr{B}$ y $\mathscr{B}''$ es*

$$[g \circ f]^{\mathscr{B}}_{\mathscr{B}''} = [g]^{\mathscr{B}'}_{\mathscr{B}''} \cdot [f]^{\mathscr{B}}_{\mathscr{B}'}.$$

Demostración. Pongamos $p = \dim U$, $n = \dim V$ y $m = \dim W$, y sean $\mathscr{B} = (x_1, \dots, x_p)$, $\mathscr{B}' = (y_1, \dots, y_n)$ y $\mathscr{B}'' = (z_1, \dots, z_m)$. Sean $f : U \to V$ y $g : V \to W$ funciones lineales y sean $[f]^{B}_{\mathscr{B}'} = (a_{i,j}) \in \mathsf{M}_{n,p}(\Bbbk)$ y $[g]^{\mathscr{B}'}_{\mathscr{B}''} = (b_{i,j}) \in \mathsf{M}_{m,n}(\Bbbk)$ sus matrices, de manera que

$$f(x_i) = a_{1,i}y_1 + \cdots + a_{n,i}y_n$$

para cada $i \in [\![p]\!]$ y

$$g(y_j) = b_{1,j}z_1 + \cdots + b_{m,j}z_m$$

para cada $j \in [\![n]\!]$. Se tiene entonces que

$$\begin{aligned}
(g \circ f)(x_i) &= g(f(x_i)) \\
&= g(a_{1,i}y_1 + \cdots + a_{n,i}y_n) \\
&= a_{1,i}g(y_1) + \cdots + a_{n,i}g(y_n) \\
&= a_{1,i}(b_{1,1}z_1 + \cdots + b_{m,1}z_m) + \cdots + a_{n,i}(b_{1,n}z_1 + \cdots + b_{m,n}z_m) \\
&= (b_{1,1}a_{1,i} + \cdots + b_{1,n}a_{n,i})z_1 + \cdots + (b_{m,1}a_{1,i} + \cdots + b_{m,n}a_{n,i})z_m
\end{aligned}$$

para cada $i \in [\![p]\!]$ y, como consecuencia de esto, que

$$\begin{aligned}
[g \circ f]^{\mathscr{B}}_{\mathscr{B}''} &= \begin{pmatrix} b_{1,1}a_{1,1} + \cdots + b_{1,n}a_{n,1} & \cdots & b_{1,1}a_{1,p} + \cdots + b_{1,n}a_{n,p} \\ \vdots & \ddots & \vdots \\ b_{m,1}a_{1,1} + \cdots + b_{m,n}a_{n,1} & \cdots & b_{m,1}a_{1,p} + \cdots + b_{m,n}a_{n,p} \end{pmatrix} \\
&= \begin{pmatrix} b_{1,1} & \cdots & b_{1,n} \\ \vdots & \ddots & \vdots \\ b_{m,1} & \cdots & b_{m,n} \end{pmatrix} \begin{pmatrix} a_{1,1} & \cdots & a_{1,p} \\ \vdots & \ddots & \vdots \\ a_{n,1} & \cdots & a_{n,p} \end{pmatrix} \\
&= [g]^{\mathscr{B}'}_{\mathscr{B}''} \cdot [f]^{\mathscr{B}}_{\mathscr{B}'}.
\end{aligned}$$

Esto es lo que queríamos probar. □

2.8.6. Proposición. *Sean V y W dos espacios vectoriales de dimensión finita, sean $n = \dim V$ y $m = \dim W$ sus dimensiones, y sean $\mathscr{B}$ y $\mathscr{B}'$ bases ordenadas de V y de W, respectivamente.*

(*i*) *Si $f : V \to W$ es un isomorfismo, entonces $m = n$, la matriz $[f]^{\mathscr{B}}_{\mathscr{B}'}$ es cuadrada e inversible, y la matriz de la función inversa f^{-1} con respecto a las bases ordenadas $\mathscr{B}'$ y $\mathscr{B}$ es*

$$[f^{-1}]^{\mathscr{B}'}_{\mathscr{B}} = \left([f]^{\mathscr{B}}_{\mathscr{B}'}\right)^{-1}.$$

(*ii*) *Recíprocamente, si $f : V \to W$ es una función lineal tal que la matriz $[f]^{\mathscr{B}}_{\mathscr{B}'}$ es inversible, entonces la función f es un isomorfismo.*

Demostración. (*i*) Como f es un isomorfismo, sabemos de la Proposición 2.4.5 que tiene que ser $m = n$. De acuerdo a la Proposición 2.8.5 y el Ejemplo 2.8.2(b) sabemos que

$$[f^{-1}]^{\mathscr{B}'}_{\mathscr{B}} \cdot [f]^{\mathscr{B}}_{\mathscr{B}'} = [f^{-1} \circ f]^{\mathscr{B}}_{\mathscr{B}} = [\mathrm{id}_V]^{\mathscr{B}}_{\mathscr{B}} = I_n$$

y que

$$[f]^{\mathscr{B}}_{\mathscr{B}'} \cdot [f^{-1}]^{\mathscr{B}'}_{\mathscr{B}} = [f \circ f^{-1}]^{\mathscr{B}'}_{\mathscr{B}'} = [\mathrm{id}_W]^{\mathscr{B}'}_{\mathscr{B}'} = I_n,$$

así que la matriz $[f]^{\mathscr{B}}_{\mathscr{B}'}$ es inversible y la matriz $[f^{-1}]^{\mathscr{B}'}_{\mathscr{B}}$ es su inversa.

(*ii*) Supongamos que $f : V \to W$ es una función lineal tal que la matriz $A = [f]^{\mathscr{B}}_{\mathscr{B}'}$ es inversible; **Hacer**: observemos que en particular esto implica que A es cuadrada, esto es, que $n = m$. De acuerdo a la Proposición 2.8.4, existe una función lineal $g : W \to V$ tal que $[g]^{\mathscr{B}'}_{\mathscr{B}} = A^{-1}$. Es

$$[g \circ f]^{\mathscr{B}}_{\mathscr{B}} = [g]^{\mathscr{B}'}_{\mathscr{B}} \cdot [f]^{\mathscr{B}}_{\mathscr{B}'} = A^{-1} \cdot A = I_n = [\mathrm{id}_V]^{\mathscr{B}}_{\mathscr{B}}$$

y, otra vez de acuerdo a la Proposición 2.8.4, esto nos dice que $g \circ f = \mathrm{id}_V$. De manera similar podemos ver que $f \circ g = \mathrm{id}_W$ y, en definitiva, que f y g son isomorfismos inversos. □

2.8.7. La matriz de una función lineal con respecto a bases ordenadas del dominio y del codominio depende, por supuesto, de la elección de esas basesr. La siguiente proposición nos dice precisamente cómo es esta dependencia.

Proposición. *Sean V y W dos espacios vectoriales de dimensión finita y sean $\mathscr{B}_1$ y $\mathscr{B}_1'$ dos bases ordenadas de V y $\mathscr{B}_2$ y $\mathscr{B}_2'$ dos bases ordenadas de W. Si $f : V \to W$ es una función lineal, entonces*

$$[f]^{\mathscr{B}_1'}_{\mathscr{B}_2'} = C(\mathscr{B}_2, \mathscr{B}_2') \cdot [f]^{\mathscr{B}_1}_{\mathscr{B}_2} \cdot C(\mathscr{B}_1', \mathscr{B}_1).$$

Demostración. Si $x \in V$, se tiene que

$$\begin{aligned} C(\mathscr{B}_2, \mathscr{B}_2') \cdot [f]^{\mathscr{B}_1}_{\mathscr{B}_2} \cdot C(\mathscr{B}_1', \mathscr{B}_1) \cdot [x]_{\mathscr{B}_1'} &= C(\mathscr{B}_2, \mathscr{B}_2') \cdot [f]^{\mathscr{B}_1}_{\mathscr{B}_2} \cdot [x]_{\mathscr{B}_1} \\ &= C(\mathscr{B}_2, \mathscr{B}_2') \cdot [f(x)]_{\mathscr{B}_2} \\ &= [f(x)]_{\mathscr{B}_2'}. \end{aligned}$$

En vista de la afirmación de unicidad de la Proposición 2.8.3, esto implica la igualdad de que aparece en el enunciado. □

§9. Proyectores

2.9.1. Proposición. *Es lo mismo tener una descomposición en suma directa que una familia de proyectores ortogonales.*

Demostración. **Hacer.** □

§10. Cocientes

2.10.1. Sea V un espacio vectorial y sea W un subespacio de V. Decimos que dos vectores x e y de V son ***congruentes módulo*** W si $x - y \in W$ y en ese caso escribimos $x \equiv y$ o, si queremos poner de manifiesto al subespacio W, $x \equiv_W y$. Esto define una relación en el conjunto V y se trata, de hecho, de una relación de equivalencia:

- Es reflexiva: si $x \in V$, entonces $x - x = 0 \in W$, así que $x \equiv x$.
- Es simétrica: si $x, y \in V$ son tales que $x \equiv y$, de manera que $x - y \in W$, entonces tenemos que $y - x = -(x - y) \in W$ y, por lo tanto, que $y \equiv x$.
- Es transitiva: si $x, y, z \in V$ son tales que $x \equiv y$ e $y \equiv z$, entonces $x - y$ e $y - z$ son elementos de W y $x - z = (x - y) + (y - z)$ también, lo que implica que $x \equiv z$.

Podemos, por lo tanto, considerar el conjunto cociente de V por esta relación de congruencia: lo escribimos V/W. Si x es un vector de V, escribiremos $[x]$ a su clase de equivalencia: así, $[x] = \{y \in V : x \equiv y\}$; si necesitamos explicitar el subespacio W, escribimos $[x]_W$. Por otro lado, si u es un elemento de V/W y x es un vector de V que está en u, de manera que $u = [x]$, decimos que x es un ***representante*** de u en V.

Queremos hacer del conjunto V/W un espacio vectorial y para ello tenemos que construir operaciones de suma y de multiplicación por escalares.

- Si x, x', y, y' son elementos de V tales que $x \equiv x'$ e $y \equiv y'$, entonces
$$(x + y) - (x' + y') = (x - x') + (y - y') \in W,$$
así que $x + y \equiv x' + y'$. Se sigue de esto que hay una función
$$+ : V/W \times V/W \to V/W$$
tal que cada vez que x e y son elementos de W se tiene que $[x] + [y] = [x + y]$.
- Si x e y son elementos de V tales que $x \equiv y$ y λ es un escalar, entonces $\lambda x \equiv \lambda y$, ya que $\lambda x - \lambda y = \lambda(x - y) \in W$. Esto implica que hay una operación
$$\cdot : \Bbbk \times V/W \to V/W$$
tal que cada vez que $\lambda \in \Bbbk$ y $x \in V$ vale que $\lambda \cdot [x] = [\lambda x]$.

2.10.2. Proposición. *Sea V un espacio vectorial y sea W un subespacio de V. El conjunto V/W dotado con las operaciones de suma y producto por escalares de $\Bbbk$ construidas arriba es un espacio vectorial. La función $\pi : x \in V \mapsto [x] \in V/W$ es una función lineal sobreyectiva cuyo núcleo es precisamente el subespacio W.*

Siempre que en la situación de esta proposición consideremos el conjunto V/W será dotado de esta estructura de espacio vectorial. Lo llamamos el ***espacio cociente de V por W***. La función π es la ***proyección canónica*** de V a V/W.

Demostración. Para ver que V/W es un espacio vectorial, tenemos que mostrar que se cumplen las condiciones de la definición 1.2.1.

- La suma es asociativa: si u, v y w son elementos de V/W, entonces existen $x, y, z \in V$ tales que $u = [x]$, $v = [y]$ y $w = [z]$, y

$$\begin{aligned}(u+v)+w &= ([x]+[y])+[z] = [x+y]+[z] = [(x+y)+z] \\ &= [x+(y+z)] = [x]+[y+z] = [x]+([y]+[z]) \\ &= u+(v+w).\end{aligned}$$

- La suma es conmutativa: si u y v son elementos de V/W, entonces existen $x, y \in V$ tales que $u = [x]$ y $v = [y]$, y se tiene que

$$u+v = [x]+[y] = [x+y] = [y+x] = [y]+[x] = v+u.$$

- La clase $[0]$ es un elemento neutro para la suma: si u es un elemento de V/W, existe $x \in V$ tal que $c = [x]$ y entonces

$$[0]+u = [0]+[x] = [0+x] = [x] = u$$

  y

$$u+x = [x]+[0] = [x+0] = [x] = u.$$

- Todo elemento de V/W posee un opuesto: si u es un elemento de V/W, entonces existe $x \in V$ tal que $u = [x]$ y

$$u+[-x] = [x]+[-x] = [x+(-x)] = [0]$$

  y

$$[-x]+u = [-x]+[x] = [(-x)+x] = [0],$$

  así que $[-x]$ es un opuesto para u.
- La multiplicación por escalares es asociativa: si $a, b \in \Bbbk$ y $u \in V/W$, entonces existe $x \in V$ tal que $u = [x]$ y

$$\begin{aligned}a\cdot(b\cdot u) &= a\cdot(b\cdot[x]) = a\cdot[b\cdot x] = [a\cdot(b\cdot x)] = [(a\cdot b)\cdot x] \\ &= (a\cdot b)\cdot[x] = (a\cdot b)\cdot u.\end{aligned}$$

- La multiplicación por escalares es unitaria: si $u \in V/W$, entonces existe $x \in V$ tal que $u = [x]$ y

$$1 \cdot u = 1 \cdot [x] = [1 \cdot x] = [x] = u.$$

- La multiplicación por escalares se distribuye sobre la suma de V/W: si $a \in \Bbbk$ y $u, v \in V/W$, entonces existen $x, y \in V$ tales que $u = [x]$ y $v = [y]$, y

$$\begin{aligned} a \cdot (u + v) &= a \cdot ([x] + [y]) = a \cdot [x + y] = [a\dot{(}x + y)] = [a \cdot x + a \cdot y] \\ &= [a \cdot x] + [a \cdot y] = a \cdot [x] + a \cdot [y] = a \cdot u + a \cdot v \end{aligned}$$

- La multiplicación por escalares se distribuye sobre la suma de $\Bbbk$: si $a, b \in \Bbbk$ y $u \in V/W$, entonces existe $x \in V$ tal que $u = [x]$ y

$$\begin{aligned} (a + b) \cdot u &= (a + b) \cdot [x] = [(a + b) \cdot x] = [a \cdot x + b \cdot x] = [a \cdot x] + [b \cdot x] \\ &= a \cdot [x] + b \cdot [x] = a \cdot u + b \cdot u \end{aligned}$$

Para ver que la función π del enunciado es una función lineal, observamos que si $a, b \in \Bbbk$ y $x, y \in V$, entonces

$$\begin{aligned} \pi(a \cdot x + b \cdot y) &= [a \cdot x + b \cdot y] = [a \cdot x] + [b \cdot y] = a \cdot [x] + b \cdot [y] \\ &= a \cdot \pi(x) + b \cdot \pi(y). \end{aligned}$$

Si $x \in V$ está en el núcleo de π, entonces su imagen $\pi(x) = [x]$ es el elemento neutro de V/W, es decir, $[x] = [0]$. Esto significa que $x \equiv 0$ y, por lo tanto, que $x = x - 0 \in W$: vemos así que $\mathsf{Nu}(\pi) \subseteq W$. Recíprocamente, si $x \in W$, entonces $x - 0 \in W$, de manera que $x \equiv 0$ y, por lo tanto, $\pi(x) = [x] = [0]$, es decir, $x \in \mathsf{Nu}(\pi)$. □

2.10.3. Una consecuencia inmediata de lo que acabamos de probar es:

Corolario. *Sea V un espacio vectorial. Todo subespacio de V es el núcleo de una función lineal de dominio V.*

Demostración. Si W es un subespacio de V, entonces la proyección $\pi : V \to V/W$ es lineal y su núcleo es W, y esto prueba el corolario. □

2.10.4. El siguiente resultado es lo que nos va a permitir en casi todos los casos construir funciones lineales cuyo dominio es un cociente:

Proposición. *Sea V un espacio vectorial, sea W un subespacio de V y sea $\pi : V \to V/W$ la proyección canónica al cociente. Si U es un espacio vectorial y $f : V \to U$ es una función lineal tal que $W \subseteq \mathsf{Nu}(f)$, entonce existe exactamente una función lineal $\bar{f} : V/W \to W$ tal que $f = \bar{f} \circ \pi$, de manera que para cada $x \in V$ es $\bar{f}([x]) = f(x)$.*

Demostración. Para cada $x, y \in V$ vale que

$$x \equiv y \implies f(x) = f(y).$$

En efecto, si $x \equiv y$ entonces $x - y \in W \subseteq \mathsf{Nu}(f)$ y por lo tanto $f(x) - f(y) = f(x - y) = 0$. Esto significa que existe exactamente una función $\bar{f} : V/W \to U$ tal que para cada $x \in V$ es $\bar{f}([x]) = f(x)$, como en el enunciado. Si $a, b \in \Bbbk$ y $u, v \in V/W$, existen $x, y \in V$ tales que $u = [x]$ y $v = [y]$, y entonces

$$\begin{aligned}\bar{f}(au + bv) &= \bar{f}(a[x] + b[y]) = \bar{f}([ax + by]) = f(ax + by) = af(x) + bf(y) \\ &= a\bar{f}([x]) + b\bar{f}([y]) = a\bar{f}(u) + b\bar{f}(v).\end{aligned}$$

Esto nos dice que la función $\bar{f}$ es lineal y completa la prueba de la proposición. □

2.10.5. Proposición. *Sea V un espacio vectorial, sea W un subespacio de V y sea $\pi : V \to V/W$ la proyección canónica al cociente.*

(*i*) *Si C es un complemento de W en V, entonces la restricción $\pi|_C : C \to V/W$ de la proyección π a C es un isomorfismo.*

(*ii*) *Si W tiene codimensión finita en V, entonces V/W tiene dimensión finita y, de hecho,*

$$\dim V/W = \operatorname{codim}_V W.$$

(*iii*) *Si V tiene dimensión finita, entonces V/W también y*

$$\dim V/W = \dim V - \dim W.$$

Demostración. Sea C un complemento de W en V y sea $\pi|_C : C \to V/W$ la restricción de π a C.

Si $u \in V/W$ y $x \in V$ es tal que $u = [x]$, entonces existen $w \in W$ y $c \in C$ con $x = w + c$ y, como $x - c = w \in W$, se tiene que $\pi(c) = [c] = [x] = u$. Esto nos dice que la restricción $\pi|_C$ es sobreyectiva. Por otro lado, si $c \in C$ es tal que $\pi|_C(c) = [c]$ es el cero $[0]$ de V/W, entonces $c - 0 \in W$, es decir, $c \in W$: como $V = W \oplus C$, vemos que $c \in W \cap C = 0$ y, en definitiva, que $\pi|_C$ es un isomorfismo. Esto prueba la primera parte de la proposición.

Para ver la segunda, supongamos que W tiene codimensión finita en W y fijemos un complemento C de W en W. De acuerdo a lo que acabamos de probar, hay un isomorfismo $C \cong V/W$ así que por un lado V/W tiene dimensión finita, porque C tiene dimensión finita, y, por otro, tenemos que $\operatorname{codim}_V W = \dim C = \dim V/W$.

Finalmente, si V tiene dimensión finita, sabemos de la Proposición 1.11.6 que el subespacio W tiene codimensión finita y que $\operatorname{codim}_V W = \dim V - \dim W$: la parte (*iii*) es entonces consecuencia inmediata de la parte (*ii*). □

2.10.6. Proposición. *Sea V un espacio vectorial, sea W un subespacio de V y sea $\pi : V \to V/W$ la proyección al cociente. Si $\mathscr{B}$ es una base de V tal que el conjunto $\mathscr{B}' = \mathscr{B} \cap W$ genera a W y ponemos $\mathscr{B}'' = \mathscr{B} \smallsetminus \mathscr{B}'$, entonces la restricción $\pi|_{\mathscr{B}''} : \mathscr{B}'' \to V/W$ es inyectiva y el conjunto $\pi(\mathscr{B}'')$ es una base de V/W.*

Demostración. De acuerdo a la Proposición 1.11.2, el conjunto $\mathscr{B}''$ es base de un complemento C de W en V, y la Proposición 2.10.5*(i)* nos dice que la restricción $\pi|_C : C \to V/W$ de la proyección π

a este complemento C es un isomorfismo. Por supuesto, se sigue de esto que la restricción $\pi|_{\mathscr{B}''}$ es inyectiva —ya que es la restricción de $\pi|_C$ al conjunto $\mathscr{B}''$— y que $\pi(\mathscr{B}'')$ es una base de V/W. □

§11. Los teoremas de isomorfismo

2.11.1. Nuestro objetivo en esta sección es probar los llamados *teoremas de isomorfismo de Noether*, que son fundamentales en la manipulación de los espacios cociente que definimos en la sección anterior. Fueron enunciados y probados originalmente —en el contexto de de los módulos y en casos especiales— por Emmy Noether (1882–1935, Alemania) en [Noe26] y, tres años más tarde —en el contexto de los grupos y con toda generalidad— por Bartel Leendert van der Waerden (1903–1996, Países Bajos) en su libro *Moderne Algebra* [vdW30], que es considerado como el tratado fundacional del álgebra moderna.

2.11.2. El *primer teorema de isomofismo* nos da una descripción a menos de isomorfismo de la imagen de una función lineal en términos de un cociente de su dominio:

Proposición. *Si $f : V \to W$ es una función lineal, entonces hay un isomorfismo*

$$\phi : V/\mathsf{Nu}(f) \to \mathsf{Im}(f)$$

tal que para cada $x \in V$ es $\phi([x]) = f(x)$.

Demostración. Sea $f : V \to W$ una función lineal. La Proposición 2.10.4 implica que existe una función lineal $\bar{f} : V/\mathsf{Nu}(f) \to \mathsf{Im}(f)$ tal que $\bar{f}([x]) = f(x)$ para todo $x \in V$. Para probar la proposición, entonces, hay que mostrar que se trata de un isomorfismo, ya que si ése es el caso podremos poner $\phi = \bar{f}$.

- Sea primero $u \in V/\mathsf{Nu}(f)$ tal que $\bar{f}(u) = 0$. Si $x \in V$ es un representante de u, de manera que $u = [x]$, entonces $\bar{f}(u) = \bar{f}([x]) = f(x) = 0$ y $x \in \mathsf{Nu}(f)$: esto nos dice que $x \equiv 0$ y, por lo tanto, que $u = [x] = [0]$ es el elemento nulo del cociente $V/\mathsf{Nu}(f)$. Así, la función $\bar{f}$ es inyectiva.
- Por otro lado, si $z \in \mathsf{Im}(f)$, existe $x \in V$ tal que $f(x) = z$ y entonces $\bar{f}([x]) = f(x) = z$: vemos así que $\bar{f}$ es sobreyectiva. □

2.11.3. El *segundo teorema de isomorfismo* afirma que ciertos subcocientes —esto es, cocientes de subespacios— de un espacio vectorial son isomorfos:

Proposición. *Sea V un espacio vectorial. Si W_1 y W_2 son dos subespacios de V, entonces hay un isomorfismo*

$$\phi : \frac{W_1}{W_1 \cap W_2} \to \frac{W_1 + W_2}{W_2}$$

tal que para cada $x \in W_1$ se tiene que $\phi([x]_{W_1 \cap W_2}) = [x]_{W_2}$.

Es usual acompañar el enunciado de esta proposición con el siguiente diagrama

$$\begin{tikzcd}[column sep=small]
 & W_1 + W_2 \arrow[dl, dash] \arrow[dr, dash] & \\
W_1 \arrow[dr, dash] & & W_2 \arrow[dl, dash] \\
 & W_1 \cap W_2 &
\end{tikzcd}$$

en el que las líneas representan las relaciones de inclusión entre los subespacios de V que aparecen en el enunciado: la conclusión de la proposición es que los cocientes correspondientes a lados opuestos de este rombo son isomorfos.

Demostración. Si $\pi : W_1 + W_2 \to (W_1 + W_2)/W_2$ la proyección canónica e $\iota : W_1 \to W_1 + W_2$ la inclusión, podemos considerar la composición

$$f = \pi \circ \iota : W_1 \to \frac{W_1 + W_2}{W_2}.$$

Ahora bien: si $x \in W_1 \cap W_2$, es

$$f(x) = \pi(\iota(x)) = \pi(x) = [x]_{W_2}$$

y, como $x \in W_2$, este último elemento del cociente $(W_1 + W_2)/W_2$ es nulo. De acuerdo a la Proposición 2.10.4, entonces, existe una función lineal $\bar{f} : W_1/(W_1 \cap W_2) \to (W_1 + W_2)/W_2$ tal que para cada $x \in W_1$ es

$$\bar{f}([x]_{W_2 \cap W_2}) = f(x) = [x]_{W_2}.$$

Mostremos que $\bar{f}$ es un isomorfismo:

- Supongamos primero que $u \in W_1/(W_1 \cap W_2)$ está en el núcleo de $\bar{f}$. Si $x \in W_1$ es un representante de u, de manera que $u = [x]_{W_1 \cap W_2}$, entonces

  $$\bar{f}(u) = f(x) = [x]_{W_2} = 0$$

  y esto significa que $x \in W_2$. Se sigue de esto que $x \in W_1 \cap W_2$ y, por lo tanto, que $u = [x]_{W_1 \cap W_2} = [0]_{W_1 \cap W_2}$, el elemento nulo de $W_1/(W_1 \cap W_2)$. Vemos así que $\bar{f}$ es inyectiva.
- Sea, por otro lado, $v \in (W_1 + W_2)/W_2$, de manera que existe $y \in W_1 + W_2$ tal que $v = [y]_{W_2}$. Esto implica que existen $y_1 \in W_1$ e $y_2 \in W_2$ con $y = y_1 + y_2$ y, en particular, que $y - y_1 = y_2 \in W_2$, esto es, que $[y]_{W_2} = [y_1]_{W_2}$. Pero entonces $\bar{f}([y_1]) = [y_1]_{W_2} = [y]_{W_2} = v$, con lo que v está en la imagen de $\bar{f}$. La función $\bar{f}$ es, por lo tanto sobreyectiva. □

2.11.4. El *tercer teorema de isomorfismo* se ocupa de cocientes interados, esto es, de describir cocientes de espacios cocientes. Para enunciarlo, necesitamos hacer antes la siguiente observación.

Si V es un espacio vectorial y W_1 y W_2 son subespacios de V tales que $W_1 \subseteq W_2$, entonces el conjunto W_2/W_1 es un subconjunto de V/W_2, esto es, toda clase de congruencia módulo W_2 en W_1 es una clase de congruencia módulo W_2 en V. Para verlo, basta observar que si $u \in W_1/W_2$ y $x \in u$, de manera que $u = [x]$, entonces

$$\{y \in W_1 : y \equiv_{W_2} x\} = \{y \in V : y \equiv_{W_2} x\}.$$

Una consecuencia inmediata de esto es que

un elemento de V/W_2 está en W_1/W_2 si y solamente si tiene un representante en V que está en W_1. (6)

Más aún, el subconjunto W_1/W_2 de V/W_2 es un subespacio: no es vacío, ya que contiene a $[0]_{W_2}$, y si $a, b \in \Bbbk$ y $u, v \in W_1/W_2$, de manera que existen $x, y \in W_1$ con $u = [x]$ y $v = [y]$, se tiene que

$$au + bv = a[x] + b[y] = [ax + by] \in W_1/W_2$$

ya que, por supuesto, $ax + by \in W_1$.

Proposición. *Sea V un espacio vectorial. Si W_1 y W_2 son subespacios de V tales que $W_1 \subseteq W_2$, entonces hay un isomorfismo*

$$\phi : \frac{V}{W_2} \to \frac{V/W_2}{W_1/W_2}$$

tal que para cada $x \in V$ vale que $\phi([x]_{W_2}) = \big[[x]_{W_1}\big]_{W_2/W_1}$.

Demostración. Sean $\pi_1 : V \to V/W_2$ y $\pi_2 : V/W_1 \to (V/W_2)/(W_1/W_2)$ las proyecciones canónicas y sea

$$f = \pi_2 \circ \pi_1 : V \to \frac{(V/W_2)}{(W_1/W_2)}$$

su composición, de manera que para cada $x \in V$ es $f(x) = \big[[x]_{W_2}\big]_{W_1/W_2}$. Si $x \in W_1$, entonces el elemento $[x]_{W_2}$ de V/W_2 pertenece a W_1/W_2, así que $\big[[x]_{W_2}\big]_{W_1/W_2}$ es el cero de $(V/W_2)/(W_1/W_2)$. La Proposición 2.10.4 nos dice por lo tanto que existe una función lineal

$$\bar{f} : \frac{V}{W_1} \to \frac{V/W_2}{W_1/W_2}$$

tal que $\bar{f}([x]_{W_1}) = \big[[x]_{W_2}\big]_{W_1/W_2}$ para todo $x \in V$. Para probar la proposición mostraremos que esta función $\bar{f}$ es un isomorfismo.

- Supongamos que $u \in V/W_1$ es tal que $\bar{f}(u) = 0$ y sea $x \in V$ un representante de u, de manera que $u = [x]_{W_1}$ y que $\bar{f}(u) = \big[[x]_{W_2}\big]_{W_1/W_2}$ es el cero de $(V/W_2)/(W_1/W_2)$. Esto nos dice que $[x]_{W_2}$ es un elemento de W_1/W_2 y, de acuerdo a la observación (6) que hicimos arriba, que $x \in W_1$: vemos así que $u = [x]_{W_1}$ es el cero de V/W_1. Esto prueba que la función $\bar{f}$ es inyectiva.

- Sea ahora v un elemento de $(V/W_2)/(W_1/W_2)$. Existe $w \in V/W_2$ tal que $v = [w]_{W_1/W_2}$ y, a su vez, existe $x \in V$ tal que $w = [x]_{W_2}$. Se sigue de esto, entonces, que

$$\bar{f}([x]_{W_1}) = \big[[x]_{W_2}\big]_{W_1/W_2} = [w]_{W_1/W_2} = v$$

y, por lo tanto, que la función $\bar{f}$ es sobreyectiva. □

2.11.5. Junto con los tres teoremas de isomorfismo anteriores es importante el siguiente *teorema de correspondencia*, que describe los subespacios de un espacio cociente.

Proposición. *Sea V un espacio vectorial y sea W un subespacio de V. Si $\pi : V \to V/W$ es la proyección canónica,*

- *$\mathscr{L}(V)_W$ es el conjunto de todos los subespacios de V que contienen a W y*
- *$\mathscr{L}(V/W)$ es el conjunto de todos los subespacios de V/W,*

entonces la función

$$\phi : U \in \mathscr{L}(V)_W \mapsto \pi(U) \in \mathscr{L}(V/W)$$

es una biyección. Más aún, si U_1 y U_2 son dos elementos de $\mathscr{L}(V)_W$, se tiene que

$$U_1 \subseteq U_2 \iff \phi(U_1) \subseteq \phi(U_2).$$

Demostración. Sea $\pi : V \to V/W$ la proyección canónica. Si U es un subespacio de V, su imagen $\pi(U)$ por π es un subespacio de V/W, así que podemos considerar la función

$$\phi : U \in \mathscr{L}(V)_W \mapsto \pi(U) \in \mathscr{L}(V/W).$$

Por otro lado, si T es un subespacio de V/W sabemos que la preimagen $\pi^{-1}(T)$ es un subespacio de V y, ya que $0 \in T$, es $W = \pi^{-1}(0) \subseteq \pi^{-1}(T)$. Esto significa que hay una función

$$\psi : T \in \mathscr{L}(V/W) \mapsto \pi^{-1}(T) \in \mathscr{L}(V)_W.$$

Afirmamos que ϕ y ψ con biyecciones inversas:

- Si $U \in \mathscr{L}(V)_W$, entonces $\psi(\phi(U)) = \pi^{-1}(\pi(U))$. Esto claramente contiene a U; por otro lado, si $x \in \pi^{-1}(\pi(U))$, entonces $\pi(x) \in \pi(U)$ y existe $u \in U$ tal que $\pi(x) = \pi(u)$: esto implica que $x - u \in \mathsf{Nu}(\pi) = W \subseteq U$ y, por lo tanto, que $x \in U$. Así, es $\psi(\phi(U)) = U$.
- Si ahora $T \in \mathscr{L}(V/W)$, entonces $\phi(\psi(T)) = \pi(\pi^{-1}(T))$. Esto está evidentemente contenido en T y, a su vez, contiene a T porque la función π es sobreyectiva. Esto nos dice que $\psi(\phi(T)) = T$.

Esto prueba la primera afirmación de la proposición. La segunda es consecuencia de que

- si $U_1, U_2 \in \mathscr{L}(V)_W$, entonces

$$U_1 \subseteq U_2 \implies \phi(U_1) = \pi(U_1) \subseteq \pi(U_2) = \phi(U_2),$$

- si $T_1, T_2 \in \mathscr{L}(V/W)$, entonces

$$T_1 \subseteq T_2 \implies \psi(T_1) = \pi^{-1}(T_1) \subseteq \pi^{-1}(T_2) = \psi(T_2)$$

y de que ϕ y ψ son biyecciones inversas. $\square$

§12. Subespacios invariantes

2.12.1. Sea V un espacio vectorial y sea $f : V \to V$ un endomorfismo de V. Un subespacio W de V es ***f-invariante*** si $f(W) \subseteq W$. Cuando ése es el caso, podemos considerar la función lineal $f_W : x \in W \mapsto f(x) \in W$ que se obtiene de f restringiendo a la vez el dominio y el codominio a W. Llamamos al endomorfismo f_W de W la ***restricción de f a W***.

Por otro lado, si W es f-invariante y $\pi : V \to V/W$ es la proyección canónica al cociente V/W, entonces $W \subseteq \mathsf{Nu}(\pi)$ y la Proposición 2.10.4 nos dice que existe exactamente una función lineal $f^W : V/W \to V/W$ tal que $f([x]) = [f(x)]$ para todo $x \in V$.

2.12.2. Proposición. *Sea V un espacio vectorial, sea $f : V \to V$ un endomorfismo de V y sea W un subespacio de V que es f-invariante.*

(*i*) *Si f_W y f^W son inyectivas, entonces f es inyectiva.*
(*ii*) *Si f_W y f^W son sobreyectivas, entonces f es sobreyectiva.*
(*iii*) *Si f es inyectiva, entonces f_W es inyectiva y, si además W tiene dimensión finita, f^W es inyectiva.*
(*iv*) *Si f es sobreyectiva, entonces f^W es sobreyectiva y, si además W tiene codimensión finita en V, f_W es sobreyectiva.*

Demostración. (*i*) Supongamos que f_W y f^W son inyectivas, y sea $x \in V$ tal que $f(x) = 0$. En ese caso tenemos que $f^W([x]) = [f(x)] = 0$ en V/W, así que la inyectividad de f^W implica que $[x] = 0$, esto es, que $x \in W$. Pero entonces $f_W(x) = f(x) = 0$ y, como f_W es inyectiva, tiene que ser $x = 0$. Esto muestra que f es inyectiva.

(*ii*) Supongamos ahora que f_W y f^W son sobreyectivas, y sea $y \in V$. Como f^W es sobreyectiva, existe $x \in V$ tal que $f^W([x]) = [f(x)] = [y]$, de manera que $y - f(x) \in W$. Como f_W es sobreyectiva, existe $z \in W$ tal que $f_W(z) = f(z) = y - f(x)$. Como entonces $y = f(x + z)$, esto nos dice que y está en la imagen de f y, en definitiva, que f es sobreyectiva.

(*iii*) Supongamos que la función f es inyectiva. Si $x \in W$ es tal que $f_W(x) = f(x) = 0$, entonces por supuesto es $x = 0$ y, por lo tanto, f_W es inyectiva.

Supongamos ahora además que W tiene dimensión finita. Como $f_W : W \to W$ es inyectiva, como acabamos de probar, es también sobreyectiva. Sea $u \in V/W$ tal que $f^W(u) = 0$. Si $x \in V$ es un representante de u en V, de manera que $u = [x]$, entonces $f^W(u) = [f(x)]$ y, como este elemento de V/W es nulo, tenemos que $f(x) \in W$. Como sabemos que f_W es sobreyectiva, esto nos dice que existe $y \in W$ tal que $f_W(y) = f(y) = f(x)$. Ahora bien, esto implica que $f(y-x) = 0$ y, como f es inyectiva, que $y - x = 0$, esto es, que $x = y \in W$. Así, es $u = [x] = [0]$. Esto prueba que f^W es inyectiva.

(iv) Finalmente, supongamos que f es sobreyectiva. Si $u \in V/W$ y $y \in V$ es un representante de u en V, entonces existe $x \in V$ tal que $f(x) = y$ y, en consecuencia, $f^W([x]) = [f(x)] = [y] = u$: vemos de esta forma que f^W es sobreyectiva.

Supongamos, para terminar, que W tiene codimensión finita. De acuerdo a la Proposición 2.10.5*(ii)*, el espacio V/W tiene dimensión finita y esto implica que el endomorfismo f^W de V/W es inyectivo, ya que es sobreyectivo. Sea $y \in W$. Como f es sobreyectiva, existe $x \in V$ tal que $f(x) = y$. Se tiene entonces que $f^W([x]) = [f(x)] = [y] = 0$, ya que $y \in W$ y, como f^W es inyectiva, que $[x] = 0$, esto es, que $x \in W$. Luego $f_W(x) = f(x) = y$ y la función f_W es sobreyectiva, como queremos □

2.12.3. Sin las hipótesis de finitud que aparecen en ellas, las últimas dos partes de la Proposición 2.12.2 dejan de ser ciertas. Veamos un ejemplo de esto: sea $\Bbbk^{\mathbb{Z}}$ es el espacio vectorial de todas las funciones $\mathbb{Z} \to \Bbbk$, sea $f : \Bbbk^{\mathbb{Z}} \to \Bbbk^{\mathbb{Z}}$ es la función lineal tal que $f(\phi)(n) = \phi(n+1)$ para cada $\phi \in \Bbbk^{\mathbb{Z}}$ y cada $n \in \mathbb{Z}$, y sea W es el subespacio de $\Bbbk^{\mathbb{Z}}$ de las funciones $\phi : \mathbb{Z} \to \Bbbk$ tales que $\phi(n) = 0$ si $n \in \mathbb{N}$. Tenemos entonces:

- El subespacio W de $\Bbbk^{\mathbb{Z}}$ es f-invariante. En efecto, si $\phi \in W$ y $n \in \mathbb{N}$, entonces
$$f(\phi)(n) = \phi(n+1) = 0,$$
ya que $n+1$ también está en $\mathbb{N}$.
- La función f es un automorfismo de $\Bbbk^{\mathbb{Z}}$ y, de hecho, su inversa es la función $g : \Bbbk^{\mathbb{Z}} \to \Bbbk^{\mathbb{Z}}$ tal que $g(\phi)(n) = \phi(n-1)$ para cada $\phi \in \Bbbk^{\mathbb{Z}}$ y cada $n \in \mathbb{Z}$.
- La función f_W no es sobreyectiva. En efecto, si $\xi : \mathbb{Z} \to \Bbbk$ es la función tal que $\xi(0) = 1$ y $\xi(n) = 0$ si $n \in \mathbb{Z} \setminus \{0\}$, entonces el único elemento ζ de $\Bbbk^{\mathbb{Z}}$ tal que $f(\zeta) = \xi$ es $g(\xi)$, pero calculándolo se ve inmediatamente que este elemento no pertenece a W.
- La función f^W no es inyectiva. Por ejemplo, si $\eta : \mathbb{Z} \to \Bbbk$ es la función tal que $\eta(1) = 1$ y $\eta(n) = 0$ para todo $n \in \mathbb{Z} \setminus \{1\}$, entonces $\eta \notin W$, de manera que $[\eta] \neq 0$ en $\Bbbk^{\mathbb{Z}}/W$, pero $f^W([\eta]) = [f(\eta)] = 0$ en ese espacio, ya que $f(\eta) \in W$.

Esto implica, claro, que W no tiene ni dimensión finita ni codimensión finita en $\Bbbk^{\mathbb{Z}}$.

Este ejemplo también muestra que no podemos eliminar la hipótesis de finitud en la segunda parte del siguiente resultado:

2.12.4. Corolario. *Sea V un espacio vectorial, sea $f : V \to V$ un endomorfismo de V y sea W un subespacio de V que es f-invariante.*

(i) *Si f_W y f^W son isomorfismos, entonces f es un isomorfismo.*
(ii) *Si f es un isomorfismo y V tiene dimensión finita, entonces f_W y f^W son isomorfismos.*

Demostración. La primera parte es consecuencia inmediata de las partes (i) y(ii) de la Proposición 2.12.2, y la segunda de las partes (iii) y (iv) de ésta, ya que si V tiene dimensión finita, entonces W tiene tanto dimensión finita como codimensión finita en V. □

2.12.5. Proposición. *Sea V un espacio vectorial de dimensión finita n y sea $f : V \to V$ un endomorfismo de V. Si W es un subespacio f-invariante de V de dimensión r y $\mathscr{B} = (x_1, \dots, x_n)$*

es una base ordenadas de V tal que $\mathscr{B}' = (x_1, \dots, x_r)$ es una base ordenada de W, entonces $\mathscr{B}'' = ([x_{r+1}], \dots, [x_n])$ es una base ordenada del cociente V/W y la matriz de f con respecto a $\mathscr{B}$ tiene una descomposición en bloques triangular superior de la forma

$$[f]_{\mathscr{B}}^{\mathscr{B}} = \begin{pmatrix} [f_W]_{\mathscr{B}'}^{\mathscr{B}'} & C \\ 0 & [f^W]_{\mathscr{B}''}^{\mathscr{B}''} \end{pmatrix}$$

con C una matriz de $\mathsf{M}_{r,n-r}(\Bbbk)$.

Demostración. **HACER.** □

Capítulo 3

Dualidad

§1. El espacio dual

3.1.1. Si V es un espacio vectorial, llamamos ***espacio dual*** a V y escribimos V^* al espacio vectorial $\hom(V, \Bbbk)$. Muchas veces llamamos a los vectores de V^* ***funcionales***.

3.1.2. Proposición. *Sea V un espacio vectorial.*

(*i*) *Si V tiene dimensión finita, entonces el espacio dual V^* también tiene dimensión finita y, de hecho, se tiene que* $\dim V = \dim V^*$.

(*ii*) *Si $n = \dim V$ y $\mathscr{B} = \{x_1, \dots, x_n\}$ es una base de V, entonces para cada $i \in [\![n]\!]$ existe exactamente una función lineal $\phi_i : V \to \Bbbk$ tal que para cada $j \in [\![n]\!]$*

$$\phi_i(x_j) = \begin{cases} 1, & \text{si } i = j; \\ 0, & \text{en caso contrario.} \end{cases} \tag{1}$$

El conjunto $\mathscr{B}^* = \{\phi_1, \dots, \phi_n\}$ es una base de V^*.

Llamamos a la base $\mathscr{B}^*$ del espacio dual V^* descripta en esta proposición la ***base dual*** a la base $\mathscr{B}$ de V.

Demostración. La primera parte es consecuencia inmediata de la Proposición 2.8.4, ya que la dimensión de $\Bbbk$ como espacio vectorial es 1, y que para cada $i \in [\![n]\!]$ existe exactamente una función lineal $\phi_i : V \to \Bbbk$ que satisface la condición (1) se sigue de la Proposición 2.1.4. Veamos, para completar la prueba, que el conjunto $\mathscr{B}^* = \{\phi_1, \dots, \phi_n\}$ es una base de V^*:

- Si $a_1, \dots, a_n \in \Bbbk$ son escalares tales que $a_1\phi_1 + \cdots + a_n\phi_n = 0$ en V^*, entonces para cada $i \in [\![n]\!]$ es

  $$0 = (a_1\phi_1 + \cdots + a_n\phi_n)(x_i) = a_1\phi_1(x_i) + \cdots + a_n\phi_n(x_i) = a_i.$$

  Esto muestra que $\mathscr{B}^*$ es un conjunto linealmente independiente.

- Por otro lado, sea $\phi \in V^*$ y consideremos el elemento $\psi = \phi(x_1)\phi_1 + \cdots + \phi(x_n)\phi_n$ de V^*. Si $i \in [\![n]\!]$, entonces

$$\psi(x_i) = (\phi(x_1)\phi_1 + \cdots + \phi(x_n)\phi_n)(x_i) = \phi(x_1)\phi_1(x_i) + \cdots + \phi(x_n)\phi_n(x_i) = \phi(x_i)$$

  y esto nos dice que ϕ y ψ son funciones lineales $V \to \Bbbk$ que coinciden sobre los elementos de la base $\mathscr{B}$: de acuerdo a la Proposición 2.1.4, entonces, es $\phi = \psi$. Como claramente ψ pertenece a $\langle \mathscr{B}^* \rangle$, vemos que $\mathscr{B}^*$ genera a V^*. □

3.1.3. Proposición. *Sea V un espacio vectorial de dimensión finita n. Si $\mathscr{B} = (x_1, \dots, x_n)$ es una base ordenada de $\mathscr{B}$ y $\mathscr{B}^* = (\phi_1, \dots, \phi_n)$ es la base de V^* dual a $\mathscr{B}$, entonces para todo $x \in V$ se tiene que*

$$x = \phi_1(x)x_1 + \cdots + \phi_n(x)x_n$$

y para todo $\phi \in V^*$ se tiene que

$$\phi = \phi(x_1)\phi_1 + \cdots + \phi(x_n)\phi_n.$$

Demostración. Sea $x \in V$. Como $\mathscr{B}$ es una base, sabemos que existen escalares $a_1, \dots, a_n \in \Bbbk$ tales que $x = a_1x_1 + \cdots + a_nx_n$. Si $i \in [\![n]\!]$, entonces

$$\phi_i(x) = \phi_i(a_1x_1 + \cdots + a_nx_n) = a_1\phi_i(x_1) + \cdots + a_n\phi_i(x_n) = a_i.$$

Esto prueba la primera afirmación. Por otro lado, la segunda afirmación fue probada en la demostración de la Proposición 3.1.2. □

3.1.4. La base dual a una base de un espacio vectorial depende claramente de ésta última: la siguiente proposición nos dice exactamente cómo.

Proposición. *Sea V un espacio vectorial de dimensión finita. Si $\mathscr{B}$ y $\mathscr{B}'$ son dos bases ordenadas de V y $\mathscr{B}^*$ y $\mathscr{B}'^*$ son las correspondientes bases duales, se tiene que*

$$C(\mathscr{B}'^*, \mathscr{B}^*) = C(\mathscr{B}, \mathscr{B}')^{\mathsf{t}}.$$

Demostración. Sea $n = \dim V$ y sean $\mathscr{B} = (x_1, \dots, x_n)$, $\mathscr{B}' = (y_1, \dots, y_n)$, $\mathscr{B}^* = (\phi_1, \dots, \phi_n)$ y $\mathscr{B}'^* = (\psi_1, \dots, \psi_n)$. Supongamos además que $C(\mathscr{B}, \mathscr{B}') = (c_{i,j})$, de manera que para cada $j \in [\![n]\!]$ se tiene que $x_j = c_{1,j}y_1 + \cdots + c_{n,j}y_n$. Si $j, k \in [\![n]\!]$, se tiene que

$$\psi_k(x_j) = \psi_k(c_{1,j}y_1 + \cdots + c_{n,j}y_n) = c_{1,j}\psi_k(y_1) + \cdots + c_{n,j}\psi_k(y_n) = c_{k,j}$$

y entonces usando la última afirmación de la Proposición 3.1.3 vemos que

$$\psi_k = \psi_k(x_1)\phi_1 + \cdots + \psi_k(x_n)\phi_n = c_{k,1}\phi_1 + \cdots + c_{k,n}\phi_n.$$

Si $C(\mathscr{B}'^*, \mathscr{B}^*) = (d_{i,j})$, esto nos dice que $d_{i,j} = c_{j,i}$ para cada $i, \in [\![n]\!]$ y prueba a proposición. □

3.1.5. Sea V un espacio vectorial. Para cada $x \in V$, hay una función $\Phi(x) : \phi \in V^* \mapsto \phi(x) \in \Bbbk$ y es fácil ver que se trata de una función lineal, de manera que $\Phi(x) \in V^{**}$, el ***espacio bidual*** de V, esto es, el espacio dual de espacio dual de V. Obtenemos de esta forma una función

$$\Phi : V \to V^{**}.$$

Proposición. *Sea V un espacio vectorial. La función $\Phi : V \to V^{**}$ que acabamos de describir es lineal. Si V tiene dimensión finita, es un isomorfismo.*

Demostración. Veamos primero que Φ es lineal: si x, $y \in V$ y s λ, $\mu \in \Bbbk$, entonces para cada funcional $\phi \in V^*$ se tiene que

$$\begin{aligned} \Phi(\lambda x + \mu y)(\phi) &= \phi(\lambda x + \mu y) \\ &= \lambda\phi(x) + \mu\phi(y) \\ &= \lambda\Phi(x)(\phi) + \mu\Phi(y)(\phi) \\ &= (\lambda\Phi(x) + \mu\Phi(y))(\phi) \end{aligned}$$

y esto significa que $\Phi(\lambda x + \mu y) = \lambda\Phi(x) + \mu\Phi(y)$.

Supongamos ahora que V tiene dimensión finita, sea $n = \dim V$, sea $\mathscr{B} = (x_1, \dots, x_n)$ una base ordenada de V y sea $\mathscr{B}^* = (\phi_1, \dots, \phi_n)$ la correspondiente base dual. Sea $x \in V$ un vector tal que $\Phi(x) = 0$. De acuerdo a la Proposición 3.1.3, sabemos que

$$\begin{aligned} x &= \phi_1(x)x_1 + \cdots + \phi_n(x)x_n \\ &= \Phi(x)(\phi_1) + \cdots + \Phi(x)(\phi_1) \\ &= 0. \end{aligned}$$

Esto muestra que la función Φ es inyectiva. Como V tiene dimensión finita, V^* también tiene dimensión finita y, de hecho, $\dim V^* = \dim V$; por la misma razón, V^{**} tiene dimensión finita y $\dim V^{**} = \dim V$. La Proposición 2.5.3 nos permite concluir entonces que Φ es un isomorfismo, como queremos. □

§2. Anuladores

Anulador a derecha

3.2.1. Sea V es un espacio vectorial y sea V^* su espacio dual. Si S es un subconjunto de V, llamamos ***anulador a derecha*** de S al conjunto

$$S^\circ = \{\phi \in V^* : \phi(x) = 0 \text{ para todo } x \in S\}.$$

3.2.2. Proposición. *Sea V un espacio vectorial y sea V^* su espacio dual.*

(i) *Se tiene que $0^\circ = V$ y $V^\circ = 0$.*

(ii) *Si S y T son subconjuntos de V y $S \subseteq T$, entonces $S^\circ \supseteq T^\circ$.*

(iii) *Si S es un subconjunto de V, entonces S° es un subespacio de V^* y coincide con $\langle S\rangle^\circ$.*

(iv) *Si S y T son subespacios de V, entonces*

$$(S+T)^\circ = S^\circ \cap T^\circ, \qquad\qquad S^\circ + T^\circ \subseteq (S\cap T)^\circ. \tag{2}$$

Si además V tiene dimensión finita, entonces esta última inclusión es una igualdad.

Demostración. (i) Todo elemento de V^* se anula en 0, así que $0^\circ = V$. Por otro lado, si un elemento de V^* se anula en todo V, necesariamente se trata de la funcón nula: esto significa que $V^\circ = 0$.

(ii) Sean S y T subconjuntos de V tales que $S \subseteq T$ y sea $\phi \in T^\circ$. Si $x \in S$, entonces $x \in T$ y, como $\phi \in T^\circ$, es $\phi(x) = 0$. Esto muestra que $\phi \in S^\circ$.

(iii) Sea S un subconjunto de V. Para cada $x \in S$ tenemos una función

$$\varepsilon_x : \phi \in V^* \mapsto \phi(x) \in \Bbbk$$

y esa función es lineal. De la definición de S° es inmediato que

$$S^\circ = \bigcap_{x\in S} \mathsf{Nu}(\varepsilon_x)$$

y entonces la Proposición 1.3.5*(ii)* nos dice que S° es un subespacio de V^*, ya que para cada $x \in S$ el subconjunto $\mathsf{Nu}(\varepsilon_x)$ lo es.

Como $S \subseteq \langle S\rangle$, la parte ($ii$) implica que $\langle S\rangle^\circ \subseteq S^\circ$. Veamos la inclusión recíproca. Sea $\phi \in S^\circ$ y sea $x \in \langle S\rangle$, de manera que existen $n \in \mathbb{N}_0$, $x_1, \dots, x_n \in S$ y $a_1, \dots, a_n \in \Bbbk$ tales que $x = a_1x_1 + \cdots + a_nx_n$: se tiene entonces que

$$\phi(x) = \phi(a_1x_1 + \cdots + a_nx_n) = a_1\phi(x_1) + \cdots + a_n\phi(x_n) = 0$$

ya que $\phi(x_i) = 0$ para todo $i \in [\![n]\!]$. Esto muestra que $x \in \langle S\rangle^\circ$ y, en definitiva, que $S^\circ S \subseteq \langle S\rangle^\circ$, como queríamos.

(iv) Sean S y T subespacios de V. Como $S \subseteq S+T$ y $T \subseteq S+T$, la parte (ii) implica que $(S+T)^\circ \subseteq S^\circ$ y $(S+T)^\circ \subseteq T^\circ$. Vemos así que, de hecho, $(S+T)^\circ \subseteq S^\circ \cap T^\circ$. Por otro lado, sea $\phi \in S^\circ \cap T^\circ$ y sea $x \in S+T$, de manera que existen $s \in S$ y $t \in T$ con $x = s+t$. Entonces $\phi(x) = \phi(s) + \phi(t) = 0$ y vemos que $\phi \in (S+T)^\circ$. Esto prueba la igualdad de (2)

Si ahora $\phi \in S^\circ + T^\circ$, de manera que existen $\sigma \in S^\circ$ y $\tau \in T^\circ$ con $\phi = \sigma + \tau$, y $x \in S \cap T$, entonces $\phi(x) = \sigma(x) + \tau(x) = 0$: esto nos dice que $\phi \in (S\cap T)^\circ$ y, en definitiva, que $S^\circ + T^\circ \subseteq (S\cap T)^\circ$.

Para terminar, supongamos que V tiene dimensión finita y mostremos que la inclusión de (2) es una igualdad. Sea $\phi \in (S\cap T)^\circ$. Como V tiene dimensión finita, sus subespacios S, T, $S\cap T$ y $S+T$ tienen todos dimensión finita. Sea $\mathscr{B} = (x_1, \dots, x_r)$ una base de $S\cap T$, sean $y_1, \dots, y_s \in V$ tales que $\mathscr{B}' = (x_1, \dots, x_t, y_1, \dots, y_s)$ es una base de S, y sean $z_1, \dots, z_t \in V$ tales que $\mathscr{B}'' = (x_1, \dots, x_r, z_1, \dots, z_t)$ es una base de T. Como en la prueba de la Proposición 1.9.4, podemos ver que $(x_1, \dots, x_r, y_1, \dots, y_s, z_1, \dots, z_t)$ es una base de $S+T$. Existen entonces $v_1, \dots, v_n \in V$

tales que $(x_1, \dots, x_r, y_1, \dots, y_s, z_1, \dots, z_t, v_1, \dots, v_n)$ es una base de V. Como consecuencia de esto, la Proposición 2.1.4 nos dice que existe una función lineal $\phi' : V \to \Bbbk$ tal que

$$\begin{aligned} &\phi'(x_i) = 0, && \text{si } i \in [\![r]\!]; && \phi'(y_i) = 0, && \text{si } i \in [\![s]\!]; \\ &\phi'(z_i) = \phi(z_i), && \text{si } i \in [\![t]\!]; && \phi'(v_i) = 0, && \text{si } i \in [\![n]\!]. \end{aligned}$$

Observemos que $\phi' \in S^\circ$, ya que esta función se anula en cada uno de los elementos de la base $\mathscr{B}'$. Sea, por otro lado, $\phi'' = \phi - \phi'$. Como

$$\phi''(x_i) = \phi(x_i) - \phi'(x_i) = 0 \qquad \text{si } i \in [\![r]\!]$$

y

$$\phi''(z_i) = \phi(z_i) - \phi'(z_i) = 0 \qquad \text{si } i \in [\![t]\!],$$

vemos que $\phi'' \in T^\circ$, y entonces $\phi = \phi' + \phi'' \in S^\circ + T^\circ$. □

3.2.3. Proposición. *Sea V un espacio vectorial de dimensión finita. Si S es un subespacio de V, entonces tanto S como su anulador a derecha S° tienen dimensión finita y*

$$\dim S + \dim S^\circ = \dim V. \tag{3}$$

Demostración. Sea S un subespacio de V. Como S y S° son subespacios de V y de V^*, respectivamente, que tienen dimensión finita, sabemos que S y S° tienen ellos también dimensión finita. Sean $n = \dim V$ y $s = \dim S$. Sea $(x_1, \dots, x_s)$ una base de S y sean $x_{s+1}, \dots, x_n \in V$ tales que $\mathscr{B} = (x_1, \dots, x_n)$ es una base de V. Sea, finalmente, $\mathscr{B}^* = (\phi_1, \dots, \phi_n)$ la base de V^* dual a $\mathscr{B}$. Para ver que vale la igualdad (3) del enunciado, es suficiente con que mostremos que $\mathscr{B}' = (\phi_{s+1}, \dots, \phi_n)$ es una base de S°. Como es un subcojunto de $\mathscr{B}$, es linealmente independiente, y sólo queda mostrar que genera a S°.

Sea $\phi \in S^\circ$. Como el conjunto $\mathscr{B}^*$ es una base de V^*, existen escalares $a_1, \dots, a_n \in \Bbbk$ tales que $\phi = a_1\phi_1 + \cdots + a_n\phi_n$. Si $i \in [\![s]\!]$, el vector x_i está en S y entonces

$$0 = \phi(x_i) = (a_1\phi_1 + \cdots + a_n\phi_n)(x_i) = a_i.$$

Esto nos dice que, de hecho, $\phi = a_{s+1}\phi_{s+1} + \cdots + a_n\phi_n \in \langle \mathscr{B}' \rangle$ y completa la prueba. □

Anulador a izquierda

3.2.4. Sea V es un espacio vectorial y sea V^* su espacio dual. Si S es un subconjunto de V^*, el ***anulador a izquierda*** de S es el conjunto

$${}^\circ S = \{x \in V : \phi(x) = 0 \text{ para todo } \phi \in S\}.$$

3.2.5. Proposición. *Sea V un espacio vectorial y sea V^* su espacio dual.*

(*i*) *Si S es un subconjunto de V^*, entonces ${}^\circ S$ es un subespacio de V.*

(*ii*) *Se tiene que ${}^\circ 0 = V$ y, si V tiene dimensión finita, que ${}^\circ(V^*) = 0$.*

(*iii*) *Si S y T son subconjuntos de V^* y $S \subseteq T$, entonces ${}^\circ S \supseteq {}^\circ T$.*

(*iv*) *Si S y T son subespacios de V^*, entonces*

$$ {}^\circ(S+T) = {}^\circ S \cap {}^\circ T, \qquad\qquad {}^\circ S + {}^\circ T \subseteq {}^\circ(S \cap T). $$

Si además V tiene dimensión finita, entonces la última inclusión es una igualdad.

Demostración. (*i*) Si $S \subseteq V^*$, entonces a definición de ${}^\circ S$ implica inmediatamente que

$$ {}^\circ S = \bigcap_{\phi \in S} \mathrm{Nu}(\phi), $$

así que ${}^\circ S$ es un subespacio por la Proposición 1.3.5*(ii)*.

(*ii*) El único elemento del subespacio nulo de V^* es la función nula y ésta que se anula en todo elemento de V: esto significa que ${}^\circ 0 = V$.

Supongamos ahora que V tiene dimensión finita y probemos que ${}^\circ(V^*) = 0$: para ello es suficiente que mostremos que

$$ \textit{para cada elemento no nulo } x \in V \setminus 0 \textit{ existe } \phi \in V^* \textit{ tal que } \phi(x) \neq 0. \tag{4} $$

Sea entonces $x \in V \setminus 0$. Como el conjunto $\{x\}$ es linealmente independiente, la Proposición 1.7.8 nos dice que si $n = \dim V$, existen vectores $x_2, \dots, x_n \in V$ tales que $(x, x_2, \dots, x_n)$ es una base de V. Si $(\phi_1, \dots, \phi_n)$ es la correspondiente base dual, entonces $\phi_1(x) = 1 \neq 0$. Esto prueba (4).

(*iii*) Sean S y T subconjuntos de V^* tales que $S \subseteq T$. Si $x \in {}^\circ T$ y $\phi \in S$, entonces $\phi(x) = 0$ porque $\phi \in T$. Vemos así que ${}^\circ T \subseteq {}^\circ S$.

(*iv*) Sean S y T subespacios de V^*. Como $S \subseteq S + T$ y $T \subseteq S + T$, la parte (*iii*) de la proposición que acabamos de probar nos dice que ${}^\circ(S+T) \subseteq {}^\circ S$ y ${}^\circ(S+T) \subseteq {}^\circ T$, de manera que ${}^\circ(S+T) \subseteq {}^\circ S \cap {}^\circ T$. Por otro lado, si $x \in {}^\circ S \cap {}^\circ T$ y $\phi \in S + T$, existen $\sigma \in S$ y $\tau \in T$ tales que $\phi = \sigma + \tau$ y, entonces, $\phi(x) = \sigma(x) + \tau(x) = 0$: esto signfica que $x \in {}^\circ(S+T)$ y, en definitiva, que ${}^\circ S \cap {}^\circ T \subseteq {}^\circ(S+T)$. Esto prueba la primera igualdad del enunciado.

Como $S \cap T \subseteq S$ y $S \cap T \subseteq T$, la parte (*iii*) nos dice que ${}^\circ S \subseteq {}^\circ(S \cap T)$ y ${}^\circ T \subseteq {}^\circ(S \cap T)$. Como ${}^\circ(S \cap T)$ es un subespacio de V, esto implica inmediatamente que ${}^\circ S + {}^\circ T \subseteq {}^\circ(S \cap T)$.

HACER: La última afirmación. □

Dualidad

3.2.6. Proposición. *Sea V un espacio vectorial.*

(*i*) *Si S es un subconjunto de V, entonces $S \subseteq {}^{\circ}(S^{\circ})$ y $S^{\circ} = ({}^{\circ}(S^{\circ}))^{\circ}$.*

(*ii*) *Si T es un subconjunto de V^*, entonces $T \subseteq ({}^{\circ}T)^{\circ}$ y ${}^{\circ}T = {}^{\circ}(({}^{\circ}T)^{\circ})$.*

Demostración. Sea S un subconjunto de V. Si $x \in S$ y $\phi \in S^{\circ}$, entonces $\phi(x) = 0$: vemos así que $x \in {}^{\circ}(S^{\circ})$ y, en definitiva, que $S \subseteq {}^{\circ}(S^{\circ})$. Esto prueba la primera afirmación de (i). La primera afirmación de (ii) se prueba de exactamente la misma forma.

Sea otra vez $S \subseteq V$ un subconjunto de V. Lo que ya probamos de (i) nos dice que $S \subseteq {}^{\circ}(S^{\circ})$ y esto implica, de acuerdo a la Proposición 3.2.2*(ii)*, que $({}^{\circ}(S^{\circ}))^{\circ} \subseteq S^{\circ}$. Por otro lado, la primera parte de (i), cuando $T = S^{\circ}$, afirma que $S^{\circ} \subseteq ({}^{\circ}(S^{\circ}))^{\circ}$. Estas dos inclusiones prueban que $S^{\circ} = ({}^{\circ}(S^{\circ}))^{\circ}$, como afirma la segunda parte de la afirmación (i). La prueba de la segunda parte de (ii) puede hacerse de manera similar. □

3.2.7. Proposición. *Sea V un espacio vectorial de dimensión finita.*

(*i*) *Si S es un subespacio de V, entonces $S = {}^{\circ}(S^{\circ})$.*

(*ii*) *Si T es un subespacio de V^*, entonces $T = ({}^{\circ}T)^{\circ}$.*

Demostración. (i) Sea S un subespacio de V. De la Proposición 3.2.6*(i)* sabemos que $S \subseteq {}^{\circ}(S^{\circ})$, así que tenemos que probar que la inclusión es una igualdad. Sea $x \in V \setminus S$. El subespacio S tiene dimensión finita. Sea $r = \dim S$ y sea $\mathscr{B} = \{x_1, \dots, x_r\}$ una base de S. Como $x \notin S$, el vector x no pertenece a $\mathscr{B}$ y el conjunto $\mathscr{B}' = \{x_1, \dots, x_r, x\}$ tiene $r + 1$ elementos y es linealmente independiente. De acuerdo a la Proposición 1.7.8, existen vectores $y_1, \dots, y_s \in V$ tales que $\mathscr{B}'' = \{x_1, \dots, x_r, x, y_1, \dots, y_s\}$ es una base de V con $r + 1 + s$ elementos. Ahora bien, existe una función lineal $\phi : V \to \Bbbk$ tal que $\phi(x_i) = 0$ para cada $i \in [\![r]\!]$, $\phi(x) = 1$ y $\phi(y_i) = 0$ para cada $i \in [\![s]\!]$. Como ϕ se anula en cada elemento de la base $\mathscr{B}$ de S, es claro que $\phi \in S^{\circ}$, y como $\phi(x) \neq 0$, además $x \notin {}^{\circ}(S^{\circ})$. Esto prueba que $V \setminus S \subseteq V \setminus {}^{\circ}(S^{\circ})$.

(ii) La Proposición 3.2.6*(ii)* nos dice que $T \subseteq ({}^{\circ}T)^{\circ}$. Para ver que vale la igualdad es suficiente entonces que mostremos que $\dim T \leq \dim({}^{\circ}T)^{\circ}$. Sean $r = \dim {}^{\circ}T$ y $n = \dim V$. De la Proposición 3.2.3 sabemos que

$$\dim({}^{\circ}T)^{\circ} = \dim V - \dim {}^{\circ}T = n - r.$$

Sea $\mathscr{B} = \{x_1, \dots, x_r\}$ una base de ${}^{\circ}T$, sean $x_{r+1}, \dots, x_n \in V$ tales que $\{x_1, \dots, x_n\}$ es una base de V, y sea $\{\phi_1, \dots, \phi_n\}$ la correspondiente base dual. Si $\phi \in T$ y $a_1, \dots, a_n \in \Bbbk$ son escalares tales que $\phi = a_1\phi_1 + \cdots + a_n\phi_n$, entonces para cada $i \in [\![r]\!]$ es

$$0 = \phi(x_i) = a_1\phi_1(x_i) + \cdots + a_n\phi_n(x_i) = a_i,$$

y entonces $\phi = a_{r+1}\phi_{r+1} + \cdots + a_n\phi_n \in \langle \phi_{r+1}, \dots, \phi_n \rangle$. Así, T está contenido en el subespacio generado por $\{\phi_{r+1}, \dots, \phi_n\}$ y, en consecuencia, tiene dimensión a lo sumo $n - r$. □

3.2.8. Proposición. *Sea V un espacio vectorial de dimensión finita. Si S es un subespacio de V^*, entonces tanto S como su anulador a izquierda ${}^\circ S$ tienen dimensión finita y*

$$\dim S + \dim {}^\circ S = \dim V.$$

Demostración. La primera afirmación es consecuencia de que S y ${}^\circ S$ son subespacios de V y de V^*, respectivamente, que tienen dimensión finita. De la Proposición 3.2.3 sabemos que

$$\dim {}^\circ S + \dim({}^\circ S)^\circ = \dim V$$

y de la Proposición 3.2.7*(ii)* que $({}^\circ S)^\circ = S$: estas dos cosas prueban la igualdad del enunciado. □

§3. Funciones transpuestas

3.3.1. Sean V y W espacios vectoriales y sea $f : V \to W$ una función lineal. Llamamos ***función transpuesta*** de f a la función

$$f^{\mathrm{t}} : \phi \in W^* \mapsto \phi \circ f \in V^*.$$

Observemos que esto tiene sentido: si $\phi : W \to \Bbbk$ es una función lineal, entonces la composición $\phi \circ f : V \to \Bbbk$ es lineal.

3.3.2. Proposición. *Sean V y W espacios vectoriales.*

(*i*) *Para cada función lineal $f : V \to W$ la función transpuesta $f^{\mathrm{t}} : V^* \to W^*$ es también lineal.*

(*ii*) *La función*

$$\tau : f \in \hom(V, W) \mapsto f^{\mathrm{t}} \in \hom(W^*, V^*)$$

es lineal.

Demostración. (*i*) Sea $f : V \to W$ una función lineal. Si $\phi, \psi \in W^*$ y $\lambda, \mu \in \Bbbk$, entonces

$$f^{\mathrm{t}}(\lambda\phi + \mu\psi) = (\lambda\phi + \mu\psi) \circ f$$

y, como la composición es una función bilineal —como afirma la Proposición 2.6.4— esto es

$$= \lambda\phi \circ f + \mu\psi \circ f = \lambda f^{\mathrm{t}}(\phi) + \mu f^{\mathrm{t}}(\psi).$$

Vemos así que la función f^{t} es lineal.

(*ii*) Sean ahora $f, g \in \hom(V, W)$ y $\lambda, \mu \in \Bbbk$. Como para cada $\phi \in W^*$ es

$$\begin{aligned}\tau(\lambda f + \mu g)(\phi) &= (\lambda f + \mu g)^{\mathrm{t}}(\phi) \\ &= \phi \circ (\lambda f + \mu g) \\ &= \lambda\phi \circ f + \mu \circ g\end{aligned}$$

$$\begin{aligned} &= \lambda f^{\mathsf{t}}(\phi) + \mu g^{\mathsf{t}}(\phi) \\ &= \lambda\tau(f)(\phi) + \mu\tau(g)(\phi) \\ &= (\lambda\tau(f) + \mu\tau(g))(\phi), \end{aligned}$$

vemos que $\tau(\lambda f + \mu g) = \lambda\tau(f) + \mu\tau(g)$, es decir, que la función τ del enunciado es lineal. □

3.3.3. Proposición.

(*i*) *Si V es un espacio vectorial e $\mathsf{id}_V : V \to V$ es la función identidad de V, entonces la función transpuesta $(\mathsf{id}_V)^t : V^* \to V^*$ es la función identidad de V^**

(*ii*) *Si U, V y W son espacios vectoriales y $f : U \to V$ y $g : V \to W$ son funciones lineales, entonces*

$$(g \circ f)^{\mathsf{t}} = f^{\mathsf{t}} \circ g^{\mathsf{t}}.$$

(*iii*) *Si V y W son espacios vectoriales y $f : V \to W$ es un isomorfismo, entonces la función transpuesta $f^{\mathsf{t}} : W^* \to V^*$ es también un isomorfismo y su función inversa es $(f^{-1})^{\mathsf{t}} : V^* \to W^*$.*

Demostración. (*i*) Sea V un espacio vectorial. Si $\phi \in V^*$, entonces

$$(\mathsf{id}_V)^{\mathsf{t}}(\phi) = \phi \circ \mathsf{id}_V = \phi_V = \mathsf{id}_{V^*}(\phi).$$

Esto significa que $(\mathsf{id}_V)^{\mathsf{t}} = \mathsf{id}_{V^*}$.

(*ii*) Sean U, V y W espacios vectoriales y sean $f : U \to V$ y $g : V \to W$ funciones lineales. Para ver que $(g \circ f)^{\mathsf{t}} = f^{\mathsf{t}} \circ g^{\mathsf{t}}$ basta observar que si $\phi \in W^*$, entonces

$$(g \circ f)^{\mathsf{t}}(\phi) = \phi \circ (g \circ f) = (\phi \circ g) \circ f = g^{\mathsf{t}}(\phi) \circ f = f^{\mathsf{t}}\big(g^{\mathsf{t}}(\phi)\big) = (g^{\mathsf{t}} \circ f^{\mathsf{t}})(\phi).$$

(*iii*) Sean V y W espacios vectoriales y sea $f : V \to W$ un isomorfismo. Para ver que $f^{\mathsf{t}} : W^* \to V^*$ es un isomorfismo con inversa $(f^{-1})^{\mathsf{t}}$ basta observar que, en vista de las partes (*ii*) y (*i*) de esta proposición, se tiene que

$$f^{\mathsf{t}} \circ (f^{-1})^{\mathsf{t}} = (f^{-1} \circ f)^{\mathsf{t}} = (\mathsf{id}_V)^{\mathsf{t}} = \mathsf{id}_{V^*}$$

y, de manera similar, que $(f^{-1} \circ f)^{\mathsf{t}} = \mathsf{id}_{W^*}$. □

3.3.4. Proposición. *Sean V y W espacios vectoriales.*

(*i*) *Si $f : V \to W$ es una función lineal entonces $\mathsf{Nu}(f^{\mathsf{t}}) = \mathsf{Im}(f)^\circ$ y $\mathsf{Im}(f^{\mathsf{t}}) \subseteq \mathsf{Nu}(f)^\circ$.*

(*ii*) *Si además V y W tienen dimensión finita, entonces esta última inclusión es una igualdad.*

Demostración. Sea $f : V \to W$ una función lineal. Si $\phi \in W^*$, es

$$\begin{aligned} \phi \in \mathsf{Nu}(f^{\mathsf{t}}) &\iff f^{\mathsf{t}}(\phi) = 0 \\ &\iff \phi \circ f = 0 \\ &\iff \text{para todo } x \in V \text{ es } \phi(f(x)) = 0 \\ &\iff \text{para todo } y \in \mathsf{Im}(f) \text{ es } \phi(y) = 0 \\ &\iff \phi \in \mathsf{Im}(f)^\circ. \end{aligned}$$

Por otro lado, si $\phi \in \mathrm{Im}(f^{\mathrm{t}})$, de manera que existe $\psi \in W^*$ tal que $\phi = f^{\mathrm{t}}(\psi) = \phi \circ f$, para cada $x \in \mathrm{Nu}(f)$ se tiene que $\phi(x) = \psi(f(x)) = 0$: esto nos dice que $\phi \in \mathrm{Nu}(f)^\circ$.

Supongamos ahora que tanto V como W tienen dimensión finita. En ese caso

$$
\begin{aligned}
\dim \mathrm{Im}(f^{\mathrm{t}}) &= \dim W^* - \dim \mathrm{Nu}(f^{\mathrm{t}}) && \text{por el Teorema 2.5.1 aplicado a } f^{\mathrm{t}} \\
&= \dim W^* - \dim \mathrm{Im}(f)^\circ && \text{por la parte } (i) \\
&= \dim W^* - (\dim W - \dim \mathrm{Im}(f)) && \text{por la Proposición 3.2.3 aplicada a } \mathrm{Im}(f) \\
&= \dim \mathrm{Im}(f) \\
&= \dim V - \dim \mathrm{Nu}(f) && \text{por el Teorema 2.5.1 aplicado a } f \\
&= \dim \mathrm{Nu}(f)^\circ && \text{por la Proposición 3.2.3 aplicada a } \mathrm{Nu}(f),
\end{aligned}
$$

y esto, junto con la inclusión $\mathrm{Im}(f^{\mathrm{t}}) \subseteq \mathrm{Nu}(f)^\circ$, prueba que vale la igualdad que queremos. □

3.3.5. Corolario. *HACER: Monos y epis, transposición.*

3.3.6. Proposición. *Sean V y W espacios vectoriales de dimensión finita y sea $f : V \to W$ una función lineal. Si $\mathscr{B}$ y $\mathscr{B}'$ son bases ordenadas de V y de W, respectivamente, y $\mathscr{B}^*$ y $\mathscr{B}'^*$ son las bases de V^* y de W^* duales a $\mathscr{B}$ y a $\mathscr{B}'$, entonces*

$$[f^{\mathrm{t}}]^{\mathscr{B}'^*}_{\mathscr{B}^*} = \left([f]^{\mathscr{B}}_{\mathscr{B}'}\right)^{\mathrm{t}}.$$

Así, la matriz de la función transpuesta de f es la matriz transpuesta de la matriz de f.

Demostración. Sean $m = \dim V$ y $n = \dim W$, y sean $\mathscr{B} = (x_1, \dots, x_m)$, $\mathscr{B}' = (y_1, \dots, y_n)$, $\mathscr{B}^* = (\phi_1, \dots, \phi_m)$ y $\mathscr{B}'^* = (\psi_1, \dots, \psi_m)$. Sea $(a_{i,j}) = [f]^{\mathscr{B}}_{\mathscr{B}'}$ la matriz de f, de manera que para cada $i \in [\![m]\!]$ es

$$f(x_i) = a_{1,i} y_1 + \cdots + a_{n,i} y_n.$$

Sea $j \in [\![n]\!]$. Como para cada $i \in [\![m]\!]$,

$$
\begin{aligned}
f^{\mathrm{t}}(\psi_j)(x_i) &= (\psi_j \circ f)(x_i) \\
&= \psi_j(f(x_i)) \\
&= \psi_j(a_{1,i} y_1 + \cdots + a_{n,i} y_n) \\
&= a_{1,i} \psi_j(y_1) + \cdots + a_{n,i} \psi_j(y_n) \\
&= a_{j,i},
\end{aligned}
$$

vemos que $f^{\mathrm{t}}(\psi_j) = a_{j,1}\phi_1 + \cdots + a_{j,m}\phi_m$. Esto significa que

$$[f^{\mathrm{t}}]^{\mathscr{B}'^*}_{\mathscr{B}^*} = \begin{pmatrix} a_{1,1} & \cdots & a_{n,1} \\ \vdots & \ddots & \vdots \\ a_{1,m} & \cdots & a_{n,m} \end{pmatrix}$$

y esta es precisamente la matriz transpuesta de $[f]^{\mathscr{B}}_{\mathscr{B}'}$. □

3.3.7. Proposición. *Sea V un espacio vectorial y sea W un subespacio de V.*

(*i*) *Si $\pi : V \to V/W$ es la proyección canónica al cociente, entonces la imagen de la función transpuesta $\pi^{\mathrm{t}} : (V/W)^* \to V^*$ está contenida en W° y la correstricción $p : (V/W)^* \to W^\circ$ de π^{t} a W° es un isomorfismo.*

(*ii*) *Si $\iota : W \to V$ es la inclusión, entonces hay un isomorfismo $q : V^*/W^\circ \to W^*$ tal que $q([\phi]) = \iota^{\mathrm{t}}(\phi)$ para toda $\phi \in V^*$.*

Demostración. **HACER.** □

3.3.8. Proposición. *Sea V un espacio vectorial, sea $f : V \to V$ un endomorfismo de V y sea W un subespacio f-invariante de V. El subespacio W° de V^* es f^{t}-invariante y si $p : (V/W)^* \to W^\circ$ y $q : V^*/W^\circ \to W^*$ son los isomorfismos de la Proposición 3.3.7, entonces conmutan los diagramas*

$$\begin{tikzcd} (V/W)^* \arrow[r, "(f^W)^{\mathrm{t}}"] \arrow[d, "p"'] & (V/W)^* \arrow[d, "p"] \\ W^\circ \arrow[r, "(f^{\mathrm{t}})_{W^\circ}"] & W^\circ \end{tikzcd} \qquad \begin{tikzcd} V^*/W^\circ \arrow[r, "(f^{\mathrm{t}})^{W^\circ}"] \arrow[d, "q"'] & V^*/W^\circ \arrow[d, "q"] \\ W^* \arrow[r, "(f_W)^{\mathrm{t}}"] & W^* \end{tikzcd}$$

Demostración. **HACER.** □

§4. El rango de una matriz

3.4.1. Si $m, n \in \mathbb{N}$ y $A \in \mathsf{M}_{m,n}(\Bbbk)$ es una matriz de m filas y n columas con entradas en $\Bbbk$, llamamos ***rango por columnas*** de A y escribimos $\mathrm{rg}(A)$ a la dimensión del subespacio de $\Bbbk^m$ generado por las n columnas de A.

3.4.2. El rango de una matriz poisee una caracterización muy sencilla:

Proposición. *Sean $m, n \in \mathbb{N}$. El rango por columnas de una matriz $A \in \mathsf{M}_{m,n}(\Bbbk)$ es igual a la dimensión de la imagen $\mathrm{Im}(f)$ de la función lineal $f : x \in \Bbbk^n \mapsto Ax \in \Bbbk^m$.*

Demostración. Sea $(e_1, \dots, e_n)$ la base ordenada estándar de $\Bbbk^n$. La imagen de la función f está generada por el conjunto $\{f(e_1), \dots, f(e_n)\}$ y, como para cada $i \in [\![n]\!]$ el vector $f(e_i) = Ae_i$ es precisamente la columna i-ésima de la matriz A, lo que afirma la proposición es inmediato. □

3.4.3. Proposición. *Sean $m, n \in \mathbb{N}$, sea $A \in \mathsf{M}_{m,n}(\Bbbk)$ y sea $f : x \in \Bbbk^n \mapsto Ax \in \Bbbk^m$.*

(*i*) *Es $0 \leq \mathrm{rg}(A) \leq \min\{m, n\}$ y el rango por columnas de A es nulo si y solamente si A es la matriz nula.*

(*ii*) *La función f es inyectiva si y solamente si $\mathrm{rg}(A) = n$.*

(*iii*) *La función f es sobreyectiva si y solamente si $\mathrm{rg}(A) = m$.*

Demostración. Como la imagen de f es un subespacio del espacio $\Bbbk^m$, que tiene dimensión m, es $\dim \operatorname{Im}(f) \leq m$. Por otro lado, del Teorema 2.5.1 es inmediato que $\dim \operatorname{Im}(f) \leq \dim \Bbbk^n = n$. Estas dos desigualdades prueban la primera afirmación de (*i*), y la segunda es evidente.

Del Teorema 2.5.2 tenemos que $n = \dim \operatorname{Nu}(f) + \operatorname{rg}(A)$, y entonces es claro que $\operatorname{Nu}(f) = 0$ si y solamente si $\operatorname{rg}(A) = n$, como afirma la parte (*ii*). Por otro lado, como $\operatorname{Im}(f)$ es un subespacio de $\Bbbk^n$, la función es sobreyectiva si y solamente si $\dim \operatorname{Im}(f) = \dim \Bbbk^n = n$: esto prueba (*iii*). □

3.4.4. Proposición. *Sean m, $n \in \mathbb{N}$ y sea $A \in \mathsf{M}_{m,n}(\Bbbk)$.*

(*i*) *Si $p \in \mathbb{N}$ y $B \in \mathsf{M}_{n,p}(\Bbbk)$, entonces*

$$\operatorname{rg}(AB) \leq \min\{\operatorname{rg}(A), \operatorname{rg}(B)\}$$

y si $\operatorname{rg}(B) = n$ entonces, de hecho,

$$\operatorname{rg}(AB) = \operatorname{rg}(A).$$

(*ii*) *Si $k \in \mathbb{N}$ y $C \in \mathsf{M}_{k,m}(\Bbbk)$ es una matriz con $\operatorname{rg}(C) = m$, entonces*

$$\operatorname{rg}(CA) = \operatorname{rg}(A).$$

Demostración. **Hacer.** □

3.4.5. El siguiente resultado es conocido como la *desigualdad de Sylvester*, por James Joseph Sylvester (1814–1897, Inglaterra):

Proposición. *Sean m, n, $p \in \mathbb{N}$. Si $A \in \mathsf{M}_{m,n}(\Bbbk)$ y $B \in \mathsf{M}_{n,p}(\Bbbk)$, entonces*

$$\operatorname{rg}(A) + \operatorname{rg}(B) - n \leq \operatorname{rg}(AB).$$

Demostración. Consideremos las funciones lineales $f : x \in \Bbbk^p \mapsto Bx \in \Bbbk^n$ y $g : x \in \Bbbk^n \mapsto Ax \in \Bbbk^m$. Del Teorema 2.5.1 y la Proposición 3.4.2 sabemos que

$$p = \dim \operatorname{Nu}(f) + \operatorname{rg}(B), \qquad n = \dim \operatorname{Nu}(g) + \operatorname{rg}(A), \qquad p = \dim \operatorname{Nu}(g \circ f) + \operatorname{rg}(AB)$$

y entonces

$$\operatorname{rg}(AB) + n = \operatorname{rg}(A) + \operatorname{rg}(B) + \big[\dim \operatorname{Nu}(f) + \dim \operatorname{Nu}(g) - \dim \operatorname{Nu}(g \circ f)\big].$$

Bastará entonces que mostremos que la expresión entre corchetes es no negativa.

Si $x \in \operatorname{Nu}(g \circ f)$ entonces claramente $f(x) \in \operatorname{Nu}(g)$, así que podemos considerar la función lineal

$$h : x \in \operatorname{Nu}(g \circ f) \mapsto f(x) \in \operatorname{Nu}(g).$$

Como $\operatorname{Im}(h) \subseteq \operatorname{Nu}(g)$, es $\dim \operatorname{Im}(h) \leq \operatorname{Nu}(g)$. Por otro lado, es claro que $\operatorname{Nu}(h) \subseteq \operatorname{Nu}(f)$, así que $\dim \operatorname{Nu}(h) \leq \dim \operatorname{Nu}(f)$. Aplicando el Teorema 2.5.1 a la función h vemos entonces que

$$\dim \operatorname{Nu}(g \circ f) = \dim \operatorname{Nu}(h) + \dim \operatorname{Im}(h) \leq \dim \operatorname{Nu}(f) + \dim \operatorname{Nu}(g),$$

y esto prueba lo que queremos. □

3.4.6. Proposición. *Sean m, $n \in \mathbb{N}$.*

(*i*) *Si A y B son elementos de $\mathsf{M}_{m,n}(\Bbbk)$, entonces*

$$\mathsf{rg}(A+B) \leq \mathsf{rg}(A) + \mathsf{rg}(B).$$

(*ii*) *Si A es un elemento de $\mathsf{M}_{m,n}(\Bbbk)$, el rango de A es el menor número $r \in \mathbb{N}_0$ tal que existen matrices $B_1, \dots, B_r \in \mathsf{M}_{m,n}(\Bbbk)$ de rango 1 tales que $A = B_1 + \cdots + B_r$.*

Demostración. (*i*) Sean A, $B \in \mathsf{M}_{m,n}(\Bbbk)$ y consideremos las funciones $f : x \in \Bbbk^n \mapsto Ax \in \Bbbk^m$ y $g : x \in \Bbbk^n \mapsto Bx \in \Bbbk^m$. Si $x \in \Bbbk^n$, entonces $(f+g)(x) = f(x) + g(x) \in \mathsf{Im}(f) + \mathsf{Im}(g)$: esto muestra que $\mathsf{Im}(f+g) \subseteq \mathsf{Im}(f) + \mathsf{Im}(g)$ y, en vista de la Proposición 1.9.4, que

$$\begin{aligned}\mathsf{rg}(A+B) &= \dim \mathsf{Im}(f+g) \\ &\leq \dim\big(\mathsf{Im}(f) + \mathsf{Im}(g)\big) \\ &= \dim \mathsf{Im}(f) + \dim \mathsf{Im}(g) - \dim\big(\mathsf{Im}(f) \cap \mathsf{Im}(g)\big) \\ &\leq \dim \mathsf{Im}(f) + \dim \mathsf{Im}(g) \\ &= \mathsf{rg}(A) + \mathsf{rg}(B).\end{aligned}$$

(*ii*) Sea $A \in \mathsf{M}_{m,n}(\Bbbk)$ y sea $r = \mathsf{rg}(A)$. Si $s \in \mathbb{N}_0$ es el menor número tal que hay matrices $B_1, \dots, B_s \in \mathsf{M}_{m,n}(\Bbbk)$ de rango 1 con $A = B_1 + \cdots + B_s$, la primera parte de la proposición implica que

$$r = \mathsf{rg}(A) = \mathsf{rg}(B_1 + \cdots + B_s) \leq \mathsf{rg}(B_1) + \cdots + \mathsf{rg}(B_s) = s.$$

Sea, por otro lado, $\{x_1, \dots, x_r\}$ una base del subespacio $\langle Ae_1, \dots, Ae_n \rangle$ de $\Bbbk^m$ generado por las columnas de A. Para cada $i \in [\![n]\!]$ existen entonces escalares $b_{i,1}, \dots, b_{i,r} \in \Bbbk$ tales que

$$Ae_i = b_{i,1}x_i + \cdots + b_{i,r}x_r.$$

Si $j \in [\![r]\!]$, sea B_j la matriz de $\mathsf{M}_{m,n}$ cuyas columnas son $b_{1,j}x_j, \dots, b_{n,j}x_j$. Esta matriz tiene rango a lo sumo 1 porque todas sus columnas son múltiplos del vector x_j. Es fácil verificar que $A = B_1 + \cdots + B_r$, así que $s \leq r$. Esto, junto con la desigualdad que obtuvimos antes, implica el resultado que buscamos. □

3.4.7. HACER: FIX THIS: De manera similar, el ***rango por filas*** de A, al que escribimos $\mathsf{rg}_F(A)$, es la dimensión del subespacio de $\Bbbk^m$ generado por las n filas de A.

Un resultado importante es que estos dos rangos son de hecho iguales:

Proposición. *Sean n, $m \in \mathbb{N}$. Si $A \in \mathsf{M}_{n,m}(\Bbbk)$, entonces* $\mathsf{rg}(A) = \mathsf{rg}_F(A)$.

En vista de esto, llamamos simplemente ***rango*** de la matriz A y escribimos $\mathsf{rg}(A)$ al valor común de $\mathsf{rg}_C(A)$ y $\mathsf{rg}_F(A)$.

Demostración. Consideremos la función lineal $f : x \in \Bbbk^m \mapsto Ax \in \Bbbk^n$. Tenemos que

$$\begin{aligned}
\dim \mathsf{Im}(f) &= n - \dim \mathsf{Im}(f)^\circ && \text{por la Proposición 3.2.3}\\
&= n - \dim \mathsf{Nu}(f^\mathsf{t}) && \text{por la Proposición 3.3.4}\\
&= n - (n - \dim \mathsf{Im}(f^\mathsf{t})) && \text{otra vez por la Proposición 3.2.3}\\
&= \dim \mathsf{Im}(f^t).
\end{aligned}$$

Si $(e_1, \dots, e_m)$ es la base estándar de $\Bbbk^m$, entonces la imagen $\mathsf{Im}(f)$ de f está generada por el conjunto $\{f(e_1), \dots, f(e_m)\}$, esto es, por las columnas de A, y entonces $\mathsf{rg}_C(A) = \dim \mathsf{Im}(f)$. Por otro lado, sabemos de la Proposición 3.3.6 que la matriz de la función transpuesta f^t con respecto a las bases duales de las bases estándares de $\Bbbk^n$ y $\Bbbk^m$ es precisamente la matriz transpuesta A^t, así que un razonamiento similar nos dice que $\mathsf{rg}_C(A^\mathsf{t}) = \dim \mathsf{Im}(f^\mathsf{t})$. Como claramente el rango por columnas de A^t coincide con el rango por filas de A, esto prueba la proposición. □

3.4.8. Nuestro último resultado sobre el rango de matrices se refiere a matrices reales y complejas:

Proposición. *Sean m, $n \in \mathbb{N}$.*

(*i*) *Si $A \in \mathsf{M}_{m,n}(\mathbb{R})$, entonces*

$$\mathsf{rg}(AA^\mathsf{t}) = \mathsf{rg}(A^\mathsf{t}A) = \mathsf{rg}(A^\mathsf{t}) = \mathsf{rg}(A).$$

(*ii*) *Si $A \in \mathsf{M}_{m,n}(\mathbb{C})$, entonces*

$$\mathsf{rg}(AA^*) = \mathsf{rg}(A^*A) = \mathsf{rg}(\overline{A}) = \mathsf{rg}(A^\mathsf{t}) = \mathsf{rg}(A^*) = \mathsf{rg}(A).$$

En la segunda parte de esta proposición, $\overline{A}$ denota a la ***matriz conjugada*** de A, esto es, a la matriz de $\mathsf{M}_{m,n}(\mathbb{C})$ cuyas entradas se obtienen conjugando a las de A, y A^* a la ***matriz adjunta*** de A, que es la matriz transpuesta de $\overline{A}$.

Demostración. **HACER.** □

Capítulo 4

Determinantes

§1. Funciones multilineales alternantes

4.1.1. Sea V un espacio vectorial. Si $d \in \mathbb{N}$, escribimos

$$V^d = \underbrace{V \times \cdots \times V}_{d \text{ factores}}$$

y decimos que una función $f : V^d \to \Bbbk$ es ***d-multilineal*** si para cada $i \in [\![d]\!]$ y cada elección de vectores $x_1, \dots, x_{i-1}, x_{i+1}, \dots, x_d$ de V la función

$$x \in V \mapsto f(x_1, \dots, x_{i-1}, x, x_{i+1}, \dots, x_d) \in \Bbbk$$

es lineal, esto es, si la función f es lineal en cada uno de sus argumentos *separadamente*. Por otro lado, si tiene la propiedad de que para toda elección de vectores $x_1, \dots, x_d \in V$ se tiene que

si existen i, $j \in [\![d]\!]$ tales que $i \neq j$ y $x_i = x_j$, entonces $f(x_1, \dots, x_d) = 0$

decimos que f es ***alternante*** y, en ese caso, decimos que el número d es el ***grado*** de f.

4.1.2. Proposición. *Sea V un espacio vectorial, sea $d \in \mathbb{N}$ y sea $f : V^d \to \Bbbk$ una función d-multilineal alternante. Si $x_1, \dots, x_d \in V$, $i \in [\![d]\!]$ e $y \in \langle x_1, \dots, \widehat{x_i}, \dots, x_d \rangle$, entonces*

$$f(x_1, \dots, x_{i-1}, x_i + y, x_{i+1}, \dots, x_d) = f(x_1, \dots, x_d).$$

Así, el valor de una función multilineal alternada no cambia si a uno de sus argumentos le sumamos una combinación lineal de los otros.

Demostración. Sean $x_1, \dots, x_d \in V$, sea $i \in \{x_1, \dots, x_d\}$ y sea $y \in \langle x_1, \dots, \widehat{x_i}, \dots, x_d \rangle$, de manera que existen escalares $a_1, \dots, a_{i-1}, a_{i+1}, \dots, a_d \in \Bbbk$ tales que

$$y = a_1 x_1 + \cdots + a_{i-1} x_{i-1} + a_{i+1} x_{i+1} + \cdots + a_d x_d.$$

Como la función f es lineal en su i-ésimo argumento, tenemos que

$$\begin{aligned}
&f(x_1,\dots,x_{i-1},y,x_{i+1},\dots,x_d)\\
&\qquad = f(x_1,\dots,x_{i-1},a_1x_1+\dots+a_{i-1}x_{i-1}+a_{i+1}x_{i+1}+\dots+a_dx_d,x_{i+1},\dots,x_d)\\
&\qquad = a_1f(x_1,\dots,x_{i-1},x_1,x_{i+1},\dots,x_d)+\dots+a_{i-1}f(x_1,\dots,x_{i-1},x_{i-1},x_{i+1},\dots,x_d)\\
&\qquad\qquad + a_{i+1}f(x_1,\dots,x_{i-1},x_{i+1},x_{i+1},\dots,x_d)+\dots+a_df(x_1,\dots,x_{i-1},x_d,x_{i+1},\dots,x_d)
\end{aligned}$$

y esta última suma es nula, ya que cada uno de sus $d-1$ términos tiene como factor a una evaluación de la función f, que es alternante, con argumentos repetidos. Usando otra vez la linealidad de f es su i-ésimo argumento, entonces, vemos que

$$\begin{aligned}
&f(x_1,\dots,x_{i-1},x_i+y,x_{i+1},\dots,x_d)\\
&\qquad = f(x_1,\dots,x_{i-1},x_i,x_{i+1},\dots,x_d)+f(x_1,\dots,x_{i-1},y,x_{i+1},\dots,x_d)\\
&\qquad = f(x_1,\dots,x_d),
\end{aligned}$$

como queremos. □

4.1.3. Proposición. *Sea V un espacio vectorial, sea $d\in\mathbb{N}$ y sea $f: V^d\to\Bbbk$ una función d-multilineal. La función f es alternante si y solamente siempre que $x_1,\ \dots,\ x_d\in V$ son vectores linealmente dependientes se tiene que $f(x_1,\dots,x_n)=0$.*

Demostración. Mostremos primer la necesidad de la condición. Supongamos que la función f es alternante y sean $x_1,\dots,x_d\in V$ vectores linealmente dependientes, de manera que existen $i\in[\![d]\!]$ y escalares $a_1,\dots,a_{i-1},a_{i+1},\dots,a_d\in\Bbbk$ tales que

$$x_i = a_1x_1+\dots+a_{i-1}x_{i-1}+a_{i+1}x_{i+1}+\dots+a_dx_d.$$

Como la función f es lineal en su i-ésimo argumento, tenemos que

$$\begin{aligned}
&f(x_1,\dots,x_{i-1},x_i,x_{i+1},\dots,x_d)\\
&\qquad = f(x_1,\dots,x_{i-1},a_1x_1+\dots+a_{i-1}x_{i-1}+a_{i+1}x_{i+1}+\dots+a_dx_d,x_{i+1},\dots,x_d)\\
&\qquad = a_1f(x_1,\dots,x_{i-1},x_1,x_{i+1},\dots,x_d)+\dots+a_{i-1}f(x_1,\dots,x_{i-1},x_{i-1},x_{i+1},\dots,x_d)\\
&\qquad\quad + a_{i+1}f(x_1,\dots,x_{i-1},x_{i+1},x_{i+1},\dots,x_d)+\dots+a_df(x_1,\dots,x_{i-1},x_d,x_{i+1},\dots,x_d).
\end{aligned}$$

Cada uno de los $d-1$ términos de esta suma tiene como factor una expresión de la forma $f(x_1,\dots,x_{i-1},x_j,x_{i+1},\dots,x_d)$ con $j\in\{1,\dots,i-1,i+1,\dots,d\}$, que se anula porque f es alternante. Esto implica, claro, que $f(x_1,\dots,x_{i-1},x_i,x_{i+1},\dots,x_d)=0$.

Veamos ahora la suficiencia. Supongamos que f satisface la condición del enunciado y sean $v_1,\dots,v_d\in V$ vectores tales que existen $i,j\in[\![d]\!]$ tales que $i\neq j$ y $x_i=x_j$. Como esto implica claramente que los vectores $x_1,\dots,x_d$ son linealmente dependientes, la hipótesis nos dice que $f(x_1,\dots,x_d)=0$. La función f es por lo tanto alternante. □

4.1.4. Sea V un espacio vectorial. Si $d \in \mathbb{N}$ y $f : V^d \to \Bbbk$ es una función multilineal, decimos que f es ***anti-simétrica*** si para cada $x_1, \dots, x_d \in V$ y cada elección de $i, j \in [\![d]\!]$ con $i < j$ se tiene que

$$f(x_1, \dots, x_i, \dots, x_j, \dots, x_d) = -f(x_1, \dots, \underbrace{x_j}_{i}, \dots, \underbrace{x_i}_{j}, \dots, x_d),$$

esto es, si al intercambiar dos de sus argumentos el valor de la función se multiplica por -1.

4.1.5. Proposición. *Sea V un espacio vectorial, sea $d \in \mathbb{N}$ y sea $f : V^d \to \Bbbk$ una función d-multilineal.*

(*i*) *Si f es alternante, entonces f es anti-simétrica.*

(*ii*) *Si f es anti-simétrica y en el cuerpo $\Bbbk$ se tiene que $2 \neq 0$, entonces f es alternante.*

Demostración. Sean $x_1, \dots, x_d \in V$ e $i, j \in [\![d]\!]$ tales que $i < j$. Como la función f es multilineal, se tiene que

$$\begin{aligned}
&f(x_1, \dots, x_{i-1}, \underbrace{x_i + x_j}_{i}, x_{i+1}, \dots, x_{j-1}, \underbrace{x_i + x_j}_{j}, x_{j+1}, \dots, x_d) \\
&\quad = f(x_1, \dots, x_{i-1}, \underbrace{x_i}_{i}, x_{i+1}, \dots, x_{j-1}, \underbrace{x_i}_{j}, x_{j+1}, \dots, x_d) \\
&\qquad + f(x_1, \dots, x_{i-1}, \underbrace{x_i}_{i}, x_{i+1}, \dots, x_{j-1}, \underbrace{x_j}_{j}, x_{j+1}, \dots, x_d) \\
&\qquad + f(x_1, \dots, x_{i-1}, \underbrace{x_j}_{i}, x_{i+1}, \dots, x_{j-1}, \underbrace{x_i}_{j}, x_{j+1}, \dots, x_d) \\
&\qquad + f(x_1, \dots, x_{i-1}, \underbrace{x_j}_{i}, x_{i+1}, \dots, x_{j-1}, \underbrace{x_j}_{j}, x_{j+1}, \dots, x_d)
\end{aligned} \tag{1}$$

Si la función f es alternante, entonces el lado izquierdo de esta igualdad y el primero y el cuarto de los sumandos del lado derecho se anulan, y entonces

$$\begin{aligned}
&f(x_1, \dots, x_{i-1}, \underbrace{x_i}_{i}, x_{i+1}, \dots, x_{j-1}, \underbrace{x_i}_{j}, x_{j+1}, \dots, x_d) \\
&\qquad\qquad + f(x_1, \dots, x_{i-1}, \underbrace{x_j}_{i}, x_{i+1}, \dots, x_{j-1}, \underbrace{x_i}_{j}, x_{j+1}, \dots, x_d) = 0.
\end{aligned}$$

Esto nos dice que en ese caso f es anti-simétrica y prueba la primera parte de la proposición.

Para ver la segunda, supongamos que f es anti-simétrica y que $2 \neq 0$ en el cuerpo $\Bbbk$. Si en la igualdad (1) de arriba es $x_i = x_j$, entonces el lado izquierdo es

$$\begin{aligned}
&f(x_1, \dots, x_{i-1}, \underbrace{2x_i}_{i}, x_{i+1}, \dots, x_{j-1}, \underbrace{2x_i}_{j}, x_{j+1}, \dots, x_d) \\
&\qquad\qquad = 2 \cdot 2 \cdot f(x_1, \dots, x_{i-1}, \underbrace{x_i}_{i}, x_{i+1}, \dots, x_{j-1}, \underbrace{x_i}_{j}, x_{j+1}, \dots, x_d)
\end{aligned}$$

mientras que el lado derecho —ya que la suma de su segundo y su tercer término es nula por la hipótesis hecha sobre f— es

$$2 \cdot f(x_1, \dots, x_{i-1}, \underbrace{x_i}_{i}, x_{i+1}, \dots, x_{j-1}, \underbrace{x_i}_{j}, x_{j+1}, \dots, x_d).$$

La igualdad, en consecuencia, nos dice que

$$\begin{aligned} 2 \cdot 2 \cdot f(x_1, \dots, x_{i-1}, \underbrace{x_i}_{i}, x_{i+1}, \dots, x_{j-1}, \underbrace{x_i}_{j}, x_{j+1}, \dots, x_d) \\ = 2 \cdot f(x_1, \dots, x_{i-1}, \underbrace{x_i}_{i}, x_{i+1}, \dots, x_{j-1}, \underbrace{x_i}_{j}, x_{j+1}, \dots, x_d). \end{aligned}$$

Por supuesto, esto implica que

$$2 \cdot f(x_1, \dots, x_{i-1}, \underbrace{x_i}_{i}, x_{i+1}, \dots, x_{j-1}, \underbrace{x_i}_{j}, x_{j+1}, \dots, x_d) = 0$$

y, como $2 \neq 0$, esto a su vez nos dice que

$$f(x_1, \dots, x_{i-1}, \underbrace{x_i}_{i}, x_{i+1}, \dots, x_{j-1}, \underbrace{x_i}_{j}, x_{j+1}, \dots, x_d) = 0.$$

Vemos así que f es alternante. □

4.1.6. Ejemplo. La condición de que 2 sea distinto de 0 en el cuerpo $\Bbbk$ que aparece en la segunda parte de la Proposición 4.1.5 es necesaria. Para ver un ejemplo, supongamos que $\Bbbk = \mathbb{F}_2$, el cuerpo de dos elementos que describimos en el Ejemplo 1.1.2(c), sea $V = \Bbbk$ y consideremos la función $f : (x, y) \in V \times V \mapsto xy \in V$ dada por la multiplicación de $\Bbbk$. Que f es bilineal es inmediato y si $x, y \in V$, entonces $f(x, y) = f(y, x)$, así que $f(x, y) + f(y, x) = 2f(x, y) = 0$. Esta función es por lo tanto bilineal y anti-simétrica. Sin embargo, no es alternante: $f(1, 1) = 1 \neq 0$. ◇

§2. Funciones multilineales alternantes de grado máximo

4.2.1. Sea V un espacio vectorial. Para cada $d \in \mathbb{N}$, escribimos $\mathrm{Alt}^d(V)$ al conjunto de todas las funciones $V^d \to \Bbbk$ que son d-multilineales y alternantes. Se trata de un subconjunto del espacio vectorial de todas las funciones $V^d \to \Bbbk$ y, de hecho, es fácil verificar que se trata de un subespacio.

4.2.2. Una consecuencia inmediata de la Proposición 4.1.3 es la siguiente:

Proposición. *Sea V un espacio vectorial. Si V tiene dimensión finita y $d > \dim V$, entonces toda función d-multilineal alternante $f : V^d \to \Bbbk$ es idénticamente nula y, en consecuencia,* $\mathrm{Alt}^d(V) = 0$.

Demostración. En efecto, si V tiene dimensión finita, $d > \dim V$ y $f : V^d \to \Bbbk$ es d-multilineal alternante, entonces cada vez que elegimos d vectores $x_1, \ldots, x_d$ en V estos son linealmente dependientes y la Proposición 4.1.3 nos dice que $f(x_1, \ldots, x_d) = 0$. □

4.2.3. Proposición. *Sea V un espacio vectorial de dimensión finita n. Si $\mathscr{B} = (e_1, \ldots, e_n)$ es una base ordenada de V, entonces la función*

$$\Phi : f \in \mathsf{Alt}^n(V) \mapsto f(e_1, \ldots, e_n) \in \Bbbk$$

es lineal e inyectiva. En particular, $\mathsf{Alt}^n(V)$ tiene dimensión finita y

$$\dim \mathsf{Alt}^n(V) \leq 1.$$

Demostración. Fijemos una base ordenada $\mathscr{B} = (e_1, \ldots, e_n)$ de V. Que la función Φ es lineal es claro. Probaremos que es inyectiva. Hecho esto, la última afirmación del enunciado seguirá inmediatamente de la Proposición 2.5.2*(ii)*.

Sea entonces $f \in \mathsf{Alt}^n(V)$ tal que $\Phi(f) = 0$ y mostremos que f es idénticamente nula. Empezamos por probar que

$$\textit{si } i_1, \ldots, i_n \in [\![n]\!], \textit{ entonces } f(e_{i_1}, \ldots, e_{i_n}) = 0. \tag{2}$$

Para ello, sean $i_1, \ldots, i_n$ elementos de $[\![n]\!]$ y para cada $k \in \{0, \ldots, n\}$ sea $P(k)$ la afirmación

$$\textit{si para cada } j \in [\![k]\!] \textit{ es } i_j = j, \textit{ entonces } f(e_{i_1}, \ldots, e_{i_n}) = 0.$$

Observemos que sabemos que $P(n)$ vale, ya que $f(e_1, \ldots, e_n) = 0$ por hipótesis, y que $P(0)$ implica que vale la afirmación (2) que queremos. Podemos entonces proceder por inducción descendente: basta que probemos que si $k \in [\![n]\!]$ y vale $P(k)$, entonces también vale $P(k-1)$.

Sea entonces $k \in [\![n]\!]$, supongamos que vale $P(k)$ y que los enteros $i_1, \ldots, i_n \in [\![n]\!]$ son tales que

$$\textit{para cada } j \in [\![k-1]\!] \textit{ se tiene que } i_j = j. \tag{3}$$

En este punto estamos en uno de los siguientes dos casos:

- Puede ser que $k \notin \{i_1, \ldots, i_n\}$. Entre los n números $i_1, \ldots, i_n$ de $[\![n]\!]$ tiene entonces que haber alguno repetido: esto es, existen $r, s \in [\![n]\!]$ tales que $r < s$ y $i_r = i_s$, y entonces $f(e_{i_1}, \ldots, e_{i_r}, \ldots, e_{i_s}, \ldots, e_{i_n}) = 0$ porque la función f es alternante.
- Puede ser, por el contrario, que $k \in \{i_1, \ldots, i_n\}$, de manera que existe $r \in [\![n]\!]$ tal que $i_r = k$. De la hipótesis (3) se sigue entonces claramente que $r \geq k$ y que

  $$f(e_{i_1}, \ldots, e_{i_{k-1}}, \underbrace{e_{i_k}}_{k}, \ldots, \underbrace{e_{i_r}}_{r}, \ldots, e_{i_n}) = -f(e_{i_1}, \ldots, e_{i_{k-1}}, \underbrace{e_{i_r}}_{k}, \ldots, \underbrace{e_{i_k}}_{r}, \ldots, e_{i_n}),$$

  ya f es anti-simétrica. Ahora bien, los primeros k argumentos de f en lado derecho de esta igualdad son, en orden, $e_1, \ldots, e_k$ y, como estamos suponiendo que vale $P(k)$, esto implica que el lado derecho es igual a 0.

En cualquiera de los dos casos tenemos entonces que $f(e_{i_1},\dots,e_{i_n}) = 0$: esto prueba que vale $P(k-1)$ y completa la inducción. Hemos verificado así que la afirmación (2) es cierta.

Queremos probar ahora que f es la función nula, esto es, que

$$\textit{si } x_1, \dots, x_n \in V, \textit{ entonces } f(x_1,\dots,x_n) = 0. \tag{4}$$

Para hacerlo, procederemos de manera similar a lo que hicimos recién: para cada $k \in \{0,\dots,n\}$ sea $Q(k)$ la afirmación

$$\textit{si } x_1, \dots, x_n \in V \textit{ y } x_1, \dots, x_k \in \mathscr{B}, \textit{ entonces } f(x_1,\dots,x_n) = 0.$$

La afirmación (4) que queremos probar es claramente equivalente a $Q(0)$ y sabemos que $Q(n)$ vale, ya que eso es precisamente lo que nos dice (2). Es suficiente entonces que, procediendo inductivamente, mostremos que si $k \in [\![k]\!]$ y vale $Q(k)$ entonces también vale $Q(k-1)$. Esto completará la prueba de la proposición.

Fijemos entonces $k \in [\![n]\!]$, supongamos que vale la afirmación $Q(k)$ y sean $x_1, \dots, x_n \in V$ tales que $x_1,\dots,x_{k-1} \in \mathscr{B}$. Como $\mathscr{B}$ es una base, hay escalares $a_1,\dots,a_n \in \Bbbk$ tales que $x_k = a_1e_1+\cdots+a_ne_n$ y, como la función f es multilineal,

$$\begin{aligned} f(x_1,\dots,\underbrace{x_k}_{k},\dots,x_n) &= f(x_1,\dots,\underbrace{a_1e_1+\cdots+a_ne_n}_{k},\dots,x_n) \\ &= a_1 f(x_1,\dots,\underbrace{e_1}_{k},\dots,x_n) + \cdots + a_n f(x_1,\dots,\underbrace{e_n}_{k},\dots,x_n). \end{aligned}$$

Cada uno de los n sumandos del último miembro de esta igualdad tiene como factor al resultado de evaluar a f en una n-upla de elementos de V cuyos primeros k elementos están en la base $\mathscr{B}$ y entonces, como estamos suponiendo que vale $Q(k)$, cada uno de ellos se anula. Así, vemos que $f(x_1,\dots,x_n) = 0$, que es lo que queríamos. □

4.2.4. Proposición. *Para cada $n \in \mathbb{N}$ existe un único elemento $D_n \in \mathsf{Alt}^n(\Bbbk^n)$ tal que*

$$D_n(e_1,\dots,e_n) = 1$$

si $(e_1,\dots,e_n)$ es la base ordenada estándar de $\Bbbk^n$.

Demostración. Es suficiente que nos ocupemos de la afirmación de existencia, ya que la unicidad se deduce de la Proposición 4.2.3. Procederemos por inducción en n, observando que como la función $D_1 : x \in \Bbbk^1 \mapsto x \in \Bbbk$ es 1-multilineal, alternante y $D_1(1) = 1$, la proposición vale si $n = 1$.

Supongamos que $n \geq 2$ y que existe una función multilineal alternante $D_{n-1} : (\Bbbk^{n-1})^{n-1} \to \Bbbk$ tal que $D_{n-1}(e_1,\dots,e_{n-1}) = 0$ si $(e_1,\dots,e_{n-1})$ es la base ordenada estándar de $\Bbbk^{n-1}$.

Si $u = (u_1,\dots,u_n)^{\mathsf{t}}$ es un elemento de $\Bbbk^n$, convengamos en escribir u' al escalar u_1 y u'' al vector $(u_2,\dots,u_n)^{\mathsf{t}}$ de $\Bbbk^{n-1}$. Es claro que si $y, z \in \Bbbk^n$ y $a, b \in \Bbbk$ se tiene que

$$(ay+bz)' = ay' + bz', \qquad\qquad (ay+bz)'' = ay'' + bz''. \tag{5}$$

Definimos una función $D_n : (\Bbbk^n)^n \to \Bbbk$ de la siguiente manera: si $x_1, \ldots, x_n \in \Bbbk^n$, ponemos

$$D_n(x_1, \ldots, x_n) = \sum_{l=1}^{n} (-1)^{l+1} x_i' D_{n-1}(x_1'', \ldots, \widehat{x_l''}, \ldots, x_n'').$$

Mostremos que D_n es n-multilineal alternante y, si $(e_1, \ldots, e_n)$ es la base ordenada estándar de $\Bbbk^n$, que $D_n(e_1, \ldots, e_n) = 1$: esto probará la proposición.

- Se tiene que $e_1' = 1$ y que $e_i' = 0$ si $i \in \{2, \ldots, n\}$, así que

$$D_n(e_1, \ldots, e_n) = \sum_{l=1}^{n} (-1)^{l+1} e_i' D_{n-1}(e_1'', \ldots, \widehat{e_l''}, \ldots, e_n'') = D_{n-1}(e_2'', \ldots, e_n'')$$

  porque en la suma todos los términos son nulos salvo posiblemente el que corresponde a $l = 1$. Como $(e_2'', \ldots, e_n'')$ es precisamente la base ordenada estándar de $\Bbbk^{n-1}$, la hipótesis inductiva nos dice que $D_n(e_1, \ldots, e_n) = 1$.
- Sea $i \in [\![n]\!]$, sean $x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_n, y, z \in \Bbbk^n$ y $a, b \in \Bbbk$. De la definición de la función D_n tenemos que

$$\begin{aligned} &D_n(x_1, \ldots, x_{i-1}, ay + bz, x_{i+1}, \ldots, x_n) \\ &\quad = \sum_{l=1}^{i-1} (-1)^{l+1} x_i' D_{n-1}(x_1'', \ldots, \widehat{x_l''}, \ldots, (ay + bz)'', \ldots, x_n'') \\ &\qquad\qquad + (-1)^{i+1} (ay + bz)' D_{n-1}(x_1'', \ldots, x_{i-1}'', x_i'', \ldots, x_n'') \\ &\qquad\qquad\qquad \sum_{l=i+1}^{n} (-1)^{l+1} x_i' D_{n-1}(x_1'', \ldots, (ay + bz)'', \ldots, \widehat{x_l''}, \ldots, x_n''). \end{aligned}$$

  Usando primero las igualdades (5) y después la multilinealidad de la función D_{n-1}, podemos reescribir esto en la forma

$$\begin{aligned} &\sum_{l=1}^{i-1} (-1)^{l+1} x_i' \Big(a D_{n-1}(x_1'', \ldots, \widehat{x_l''}, \ldots, y'', \ldots, x_n'') \\ &\qquad\qquad\qquad + b D_{n-1}(x_1'', \ldots, \widehat{x_l''}, \ldots, z'', \ldots, x_n'') \Big) \\ &\quad + (-1)^{i+1} a y' D_{n-1}(x_1'', \ldots, x_{i-1}'', x_i'', \ldots, x_n'') \\ &\qquad\qquad\qquad + (-1)^{i+1} b z' D_{n-1}(x_1'', \ldots, x_{i-1}'', x_i'', \ldots, x_n'') \\ &\quad + \sum_{l=i+1}^{n} (-1)^{l+1} x_i' \Big(a D_{n-1}(x_1'', \ldots, y'', \ldots, \widehat{x_l''}, \ldots, x_n'') \\ &\qquad\qquad\qquad + b D_{n-1}(x_1'', \ldots, z'', \ldots, \widehat{x_l''}, \ldots, x_n'') \Big). \end{aligned}$$

  Por otro lado,

$$a D_n(x_1, \ldots, x_{i-1}, y, x_{i+1}, \ldots, x_n) + b D_n(x_1, \ldots, x_{i-1}, z, x_{i+1}, \ldots, x_n)$$

$$
\begin{aligned}
&= a\sum_{l=1}^{i-1}(-1)^{l+1}x'_iD_{n-1}(x''_1,\dots,\widehat{x''_l},\dots,y'',\dots,x''_n)\\
&\qquad\qquad + a(-1)^{i+1}y'D_{n-1}(x''_1,\dots,x''_{i-1},x''_i,\dots,x''_n)\\
&\qquad\qquad\qquad a\sum_{l=i+1}^{n}(-1)^{l+1}x'_iD_{n-1}(x''_1,\dots,y'',\dots,\widehat{x''_l},\dots,x''_n)\\
&\quad + b\sum_{l=1}^{i-1}(-1)^{l+1}x'_iD_{n-1}(x''_1,\dots,\widehat{x''_l},\dots,z'',\dots,x''_n)\\
&\qquad\qquad + b(-1)^{i+1}y'D_{n-1}(x''_1,\dots,x''_{i-1},x''_i,\dots,x''_n)\\
&\qquad\qquad\qquad b\sum_{l=i+1}^{n}(-1)^{l+1}x'_iD_{n-1}(x''_1,\dots,z'',\dots,\widehat{x''_l},\dots,x''_n)
\end{aligned}
$$

y esto es igual a la expresión anterior. Vemos así que la función D_n es lineal con respecto a su i-ésimo argumento.

- Sean finalmente $x_1,\dots,x_n \in V$ y supongamos que $r, s \in [\![n]\!]$ son tales que $r < s$ y $x_r = x_s$. De la definición de D_n tenemos que

$$D_n(x_1,\dots,x_n) = \sum_{l=1}^{n}(-1)^{l+1}x'_lD_{n-1}(x''_1,\dots,\widehat{x''_l},\dots,x''_n).$$

Si el índice $l \in [\![n]\!]$ es distinto de r y de s, el término l-ésimo de esta suma tiene a $D_{n-1}(x''_1,\dots,\widehat{x''_l},\dots,x''_n)$ como factor, y este escalar es nulo porque D_{n-1} es una función alternante: entre sus argumentos aparecen los vectores x''_r y x''_s, que son iguales. Esto significa que, de hecho,

$$
\begin{aligned}
&D_n(x_1,\dots,x_n)\\
&\quad = (-1)^{r+1}x'_rD_{n-1}(x''_1,\dots,\widehat{x''_r},\dots,x''_n) + (-1)^{s+1}x'_sD_{n-1}(x''_1,\dots,\widehat{x''_s},\dots,x''_n).
\end{aligned}
\tag{6}
$$

Como $x_r = x_s$, tenemos por supuesto que $x'_r = x'_s$ y $x''_r = x''_s$. Como podemos pasar de la secuencia de $n-1$ vectores

$$x''_1,\dots,\widehat{x''_r},\dots,x''_s,\dots,x''_n$$

a la secuencia

$$x''_1,\dots,x''_r,\dots,\widehat{x''_s},\dots,x''_n$$

haciendo $s-r-1$ transposiciones —la Figura 4.1 de la página siguiente ilustra esto— el hecho de que D_{n-1} sea alternante y, por lo tanto, anti-simétrica nos dice que

$$D_{n-1}(x''_1,\dots,\widehat{x''_s},\dots,x''_n) = (-1)^{s-r-1}D_{n-1}(x''_1,\dots,\widehat{x''_r},\dots,x''_n)$$

y entonces, reemplazando en la igualdad (6), vemos que

$$D_n(x_1,\dots,x_n) = \big((-1)^{r+1} + (-1)^{s+1}(-1)^{s-r-1}\big)\cdot x'_rD_{n-1}(x''_1,\dots,\widehat{x''_r},\dots,x''_n)$$

y esto es igual a 0, porque el primer factor es nulo. La proposición queda así probada. □

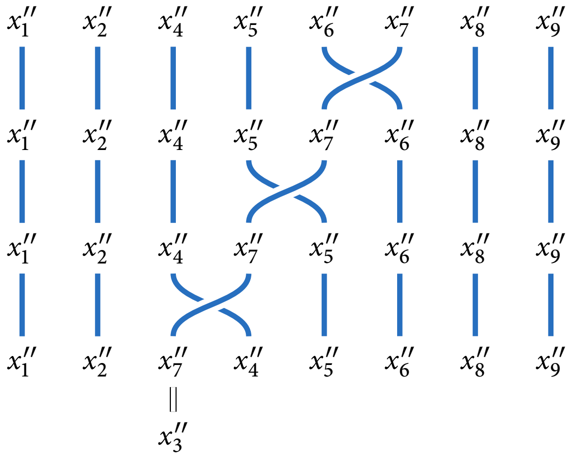


Figura 4.1. El caso en que $n = 9$, $r = 3$ y $s = 7$: si $x_3'' = x_7''$, para pasar de $(x_1'', \dots, \widehat{x_3''}, \dots, x_9'')$ a $(x_1'', \dots, \widehat{x_7''}, \dots, x_9'')$ son necesarias $3 = s - r - 1$ transposiciones.

§3. El determinante de una matriz

4.3.1. Sea $n \in \mathbb{N}$ y sea $(e_1, \dots, e_n)$ la base ordenada estándar de $\Bbbk^n$. Si $A \in \mathsf{M}_n(\Bbbk)$ es una matriz cuadrada de n filas y n columnas con entradas en $\Bbbk$, el ***determinante*** de A, al que escribimos $\det A$ o $|A|$, es el escalar $D_n(Ae_1, \dots, Ae_n)$ que se obtiene de evaluar la función $D_n : (\Bbbk^n)^n \to \Bbbk$ de la Proposición 4.2.4 en los n vectores $Ae_1, \dots, Ae_n$ de $\Bbbk^n$. Observemos que estos vectores son precisamente las columnas de la matriz A. Si A es la matriz

$$\begin{pmatrix} a_{1,1} & \cdots & a_{1,n} \\ \vdots & \ddots & \vdots \\ a_{n,1} & \cdots & a_{n,n} \end{pmatrix}$$

escribimos muchas veces al determinante de A usando la notación

$$\begin{vmatrix} a_{1,1} & \cdots & a_{1,n} \\ \vdots & \ddots & \vdots \\ a_{n,1} & \cdots & a_{n,n} \end{vmatrix}.$$

La prueba de la Proposición 4.2.4 da una construcción recursiva de la función D_n, que la expresa en términos de D_{n-1}: siguiendo esa construcción, vemos que

- Si $n = 1$, entonces

$$|a_{1,1}| = D_1((a_{1,1})) = a_{1,1}.$$

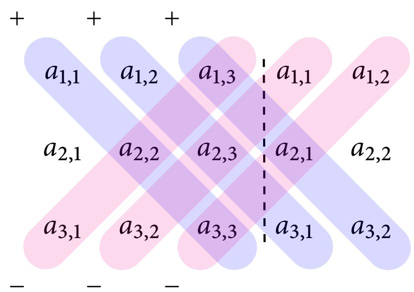
- Si $n = 2$, entonces

$$\begin{aligned}\begin{vmatrix} a_{1,1} & a_{1,2} \\ a_{2,1} & a_{2,2} \end{vmatrix} &= D_2\left(\begin{pmatrix} a_{1,1} \\ a_{2,1} \end{pmatrix}, \begin{pmatrix} a_{1,2} \\ a_{2,2} \end{pmatrix}\right) \\ &= a_{1,1}D_1((a_{2,2})) - a_{1,2}D_1((a_{2,1})) \\ &= a_{1,1}a_{2,2} - a_{1,2}a_{2,1}.\end{aligned}$$

- Si $n = 3$, entonces

$$\begin{aligned}\begin{vmatrix} a_{1,1} & a_{1,2} & a_{1,3} \\ a_{2,1} & a_{2,2} & a_{2,3} \\ a_{3,1} & a_{3,2} & a_{3,3} \end{vmatrix} &= D_3\left(\begin{pmatrix} a_{1,1} \\ a_{2,1} \\ a_{3,1} \end{pmatrix}, \begin{pmatrix} a_{1,2} \\ a_{2,2} \\ a_{3,2} \end{pmatrix}, \begin{pmatrix} a_{1,3} \\ a_{2,3} \\ a_{3,3} \end{pmatrix}\right) \\ &= a_{1,1}D_2\left(\begin{pmatrix} a_{2,2} \\ a_{3,2} \end{pmatrix}, \begin{pmatrix} a_{2,3} \\ a_{3,3} \end{pmatrix}\right) - a_{1,2}D_2\left(\begin{pmatrix} a_{2,1} \\ a_{3,1} \end{pmatrix}, \begin{pmatrix} a_{2,3} \\ a_{3,3} \end{pmatrix}\right) + a_{1,3}D_2\left(\begin{pmatrix} a_{2,1} \\ a_{3,1} \end{pmatrix}, \begin{pmatrix} a_{2,2} \\ a_{3,2} \end{pmatrix}\right) \\ &= a_{1,1}a_{2,2}a_{3,3} - a_{1,1}a_{3,2}a_{2,3} - a_{1,2}a_{2,1}a_{3,3} + a_{1,2}a_{3,1}a_{2,3} + a_{1,3}a_{2,1}a_{3,2} - a_{1,3}a_{3,1}a_{2,2}.\end{aligned}$$

Una forma de recordar esta última expresión es la siguiente *regla de Sarrus*, por Pierre Frédéric Sarrus (1798–1861, Francia): si copiamos las dos primeras columnas a la derecha de la matriz, para obtener una matriz de 3 filas y 5 columnas,



entonces el determinante de la matriz es la suma de los tres productos correspondientes a las diagonales azules del dibujo menos la suma de los tres productos correspondientes a las rojas.

- Si $n = 4$, entonces

$$\begin{vmatrix} a_{1,1} & a_{1,2} & a_{1,3} & a_{1,4} \\ a_{2,1} & a_{2,2} & a_{2,3} & a_{2,4} \\ a_{3,1} & a_{3,2} & a_{3,3} & a_{3,4} \\ a_{4,1} & a_{4,2} & a_{4,3} & a_{4,4} \end{vmatrix} = D_4\left(\begin{pmatrix} a_{1,1} \\ a_{2,1} \\ a_{3,1} \\ a_{4,1} \end{pmatrix}, \begin{pmatrix} a_{1,2} \\ a_{2,2} \\ a_{3,2} \\ a_{4,2} \end{pmatrix}, \begin{pmatrix} a_{1,3} \\ a_{2,3} \\ a_{3,3} \\ a_{4,3} \end{pmatrix}, \begin{pmatrix} a_{1,4} \\ a_{2,4} \\ a_{3,4} \\ a_{4,4} \end{pmatrix}\right)$$

$$
\begin{aligned}
&= a_{1,1}D_3\left(\begin{pmatrix}a_{2,2}\\a_{3,2}\\a_{4,2}\end{pmatrix},\begin{pmatrix}a_{2,3}\\a_{3,3}\\a_{4,3}\end{pmatrix},\begin{pmatrix}a_{2,4}\\a_{3,4}\\a_{4,4}\end{pmatrix}\right) - a_{1,2}D_3\left(\begin{pmatrix}a_{2,1}\\a_{3,1}\\a_{4,1}\end{pmatrix},\begin{pmatrix}a_{2,3}\\a_{3,3}\\a_{4,3}\end{pmatrix},\begin{pmatrix}a_{2,4}\\a_{3,4}\\a_{4,4}\end{pmatrix}\right)\\
&\qquad + a_{1,3}D_3\left(\begin{pmatrix}a_{2,1}\\a_{3,1}\\a_{4,1}\end{pmatrix},\begin{pmatrix}a_{2,2}\\a_{3,2}\\a_{4,2}\end{pmatrix},\begin{pmatrix}a_{2,4}\\a_{3,4}\\a_{4,4}\end{pmatrix}\right) - a_{1,4}D_3\left(\begin{pmatrix}a_{2,1}\\a_{3,1}\\a_{4,1}\end{pmatrix},\begin{pmatrix}a_{2,2}\\a_{3,2}\\a_{4,2}\end{pmatrix},\begin{pmatrix}a_{2,3}\\a_{3,3}\\a_{4,3}\end{pmatrix},\right)\\
&= \quad a_{1,1}, a_{2,2}, a_{3,3}, a_{4,4} - a_{1,1}, a_{2,2}, a_{4,3}, a_{3,4} - a_{1,1}, a_{3,2}, a_{2,3}, a_{4,4}\\
&\quad +a_{1,1}, a_{3,2}, a_{4,3}, a_{2,4} + a_{1,1}, a_{4,2}, a_{2,3}, a_{3,4} - a_{1,1}, a_{4,2}, a_{3,3}, a_{2,4}\\
&\quad -a_{2,1}, a_{1,2}, a_{3,3}, a_{4,4} + a_{2,1}, a_{1,2}, a_{4,3}, a_{3,4} + a_{2,1}, a_{3,2}, a_{1,3}, a_{4,4}\\
&\quad -a_{2,1}, a_{3,2}, a_{4,3}, a_{1,4} - a_{2,1}, a_{4,2}, a_{1,3}, a_{3,4} + a_{2,1}, a_{4,2}, a_{3,3}, a_{1,4}\\
&\quad +a_{3,1}, a_{1,2}, a_{2,3}, a_{4,4} - a_{3,1}, a_{1,2}, a_{4,3}, a_{2,4} - a_{3,1}, a_{2,2}, a_{1,3}, a_{4,4}\\
&\quad +a_{3,1}, a_{2,2}, a_{4,3}, a_{1,4} + a_{3,1}, a_{4,2}, a_{1,3}, a_{2,4} - a_{3,1}, a_{4,2}, a_{2,3}, a_{1,4}\\
&\quad -a_{4,1}, a_{1,2}, a_{2,3}, a_{3,4} + a_{4,1}, a_{1,2}, a_{3,3}, a_{2,4} + a_{4,1}, a_{2,2}, a_{1,3}, a_{3,4}\\
&\quad -a_{4,1}, a_{2,2}, a_{3,3}, a_{1,4} - a_{4,1}, a_{3,2}, a_{1,3}, a_{2,4} + a_{4,1}, a_{3,2}, a_{2,3}, a_{1,4}.
\end{aligned}
$$

Podemos continuar de esta forma, pero rápidamente el tamaño de las fórmulas crece: puede verse en las que acabamos de dar que la cantidad de términos en el determinante de una matriz cuadrada de tamaño 1, 2, 3 y 4 es, respectivamente, 1, 2, 6 y 24. Más generalmente, es fácil deducir inductivamente de la prueba de la Proposición 4.2.4 que tenemos una expresión con a lo sumo $n!$ términos para el determinante de una matriz cuadrada de tamaño n.

4.3.2. Proposición. *Sea $n \in \mathbb{N}$. La función* $\det : \mathsf{M}_n(\Bbbk) \to \Bbbk$ *tiene las siguientes propiedades:*

(*i*) $\det A$ *depende multilinealmente de las columnas de la matriz A.*

(*ii*) *Si $A \in \mathsf{M}_n(\Bbbk)$ tiene dos columnas iguales, entonces* $\det A = 0$.

(*iii*) *Si A y B son elementos de $\mathsf{M}_n(\Bbbk)$ y B se obtiene de A intercambiando dos columnas, entonces* $\det B = -\det A$.

(*iv*) *Si A y B son elementos de $\mathsf{M}_n(\Bbbk)$ y B se obtiene de A sumando a una de sus columnas una combinación lineal de las demás, entonces* $\det B = \det A$.

Demostración. Todo esto es claro, en vista de la definición de det y de la Proposición 4.2.4. □

4.3.3. Proposición. *Sea $n \in \mathbb{N}$.*

(*i*) *Si $I_n \in \mathsf{M}_n(\Bbbk)$ es la matriz identidad, entonces* $\det I_n = 1$.

(*ii*) *Si A y B son matrices de $\mathsf{M}_n(\Bbbk)$, entonces* $\det AB = \det A \cdot \det B$.

(*iii*) *Si $A \in \mathsf{M}_n(\Bbbk)$ es inversible, entonces* $\det A \neq 0$ *y, de hecho,* $\det A^{-1} = (\det A)^{-1}$.

Demostración. (*i*) De acuerdo a nuestras definiciones y la elección de la función D_n, es

$$\det I_n = D_n(I_n e_1, \dots, I_n e_n) = D_n(e_1, \dots, e_n) = 1.$$

(*ii*) Sean A y B dos elementos de $\mathsf{M}_n(\Bbbk)$ y sea $\tilde{D} : (\Bbbk^n)^n \to \Bbbk$ la función tal que cada vez que

$x_1, \dots, x_n$ son elementos de $\Bbbk^n$ es

$$\tilde{D}(x_1, \dots, x_n) = D_n(Ax_1, \dots, Ax_n).$$

Como $D_n : (\Bbbk^n)^n \to \Bbbk$ es una función n-multilineal y alternante, es inmediato verificar que lo mismo es cierto de $\tilde{D}$, esto es, que $\tilde{D}$ es un elemento del espacio vectorial $\mathsf{Alt}^n(\Bbbk^n)$. Sabemos que este espacio vectorial tiene dimensión 1 y que $\{D_n\}$ es una de sus bases: esto significa que existe un escalar $\alpha \in \Bbbk$ tal que $\tilde{D} = \alpha D_n$ y, de hecho, si $(e_1, \dots, e_n)$ es la base ordenada estándar de $\Bbbk^n$ tenemos que

$$\alpha = \alpha D_n(e_1, \dots, e_n) = \tilde{D}(e_1, \dots, e_n) = D_n(Ae_1, \dots, Ae_n) = \det A.$$

Esto implica que

$$\tilde{D}(Be_1, \dots, Be_n) = \det A \cdot D(Be_1, \dots, Be_n) = \det A \cdot \det B$$

mientras que la definición misma de la función $\tilde{D}$ nos dice, por otro lado, que

$$\tilde{D}(Be_1, \dots, Be_n) = D_n(ABe_1, \dots, ABe_n) = \det AB.$$

Esto prueba la afirmación (*ii*) de la proposición.

(*iii*) Sea $A \in \mathsf{M}_n(\Bbbk)$ una matriz inversible. Usando las partes (*ii*) y (*i*) de la proposición, vemos que

$$\det A \cdot \det A^{-1} = \det AA^{-1} = \det I_n = 1,$$

así que claramente $\det A \neq 0$ y $\det A^{-1} = (\det A)^{-1}$, que es lo que queríamos probar. □

4.3.4. Un corolario inmediato de la Proposición 4.3.3 pero importante es el siguiente:

Corolario. *Si A y B son elementos de $\mathsf{M}_n(\Bbbk)$ para los que existe una matriz $C \in \mathsf{M}_n(\Bbbk)$ inversible tal que $A = CBC^{-1}$, entonces $\det A = \det B$.*

Demostración. En efecto, en ese caso la Proposición 4.3.3 nos dice que $\det C \neq 0$ y que

$$\det A = \det CBC^{-1} = \det C\dot{\det}B \cdot C^{-1} = \det C\dot{\det}B(\cdot C)^{-1} = \det B,$$

como afirma el corolario. □

§4. La fórmula de Leibniz

4.4.1. Si $n \in \mathbb{N}$, escribamos $[\![n]\!]$ al conjunto $[\![n]\!]$. Una ***permutación*** de $[\![n]\!]$ es una función biyectiva $\sigma : [\![n]\!] \to [\![n]\!]$ y S_n es el conjunto de todas las permutaciones de $[\![n]\!]$.

Proposición. *Sea $n \in \mathbb{N}$.*

(i) *El conjunto* S_n *tiene exactamente* $n!$ *elementos.*
(ii) *La función identidad* $\mathrm{id} : [\![n]\!] \to [\![n]\!]$ *es claramente una permutación.*
(iii) *Si* σ *y* τ *son dos permutaciones de* $[\![n]\!]$, *entonces la composición* $\sigma \circ \tau$ *también lo es.*
(iv) *Si* σ *es una permutación de* $[\![n]\!]$, *entonces la función inversa* σ^{-1} *—que existe porque* σ *es biyectiva— también es una permutación de* $[\![n]\!]$.

Demostración. Todas las afirmaciones son inmediatas. □

4.4.2. Cuando n es pequeño, escribimos a veces a una permutación $\sigma \in S_n$ en la forma de una matriz de 2 filas y n columnas

$$\begin{pmatrix} 1 & 2 & \cdots & n \\ \sigma(1) & \sigma(2) & \cdots & \sigma(n) \end{pmatrix}$$

listando en la primera fila los elementos de $[\![n]\!]$ y en la segunda las correspondientes imágenes por σ. De manera todavía más compacta, escribimos a σ también como la secuencia ordenada de sus valores

$$\sigma(1)\,\sigma(2)\cdots\sigma(n).$$

Así, por 35142 y por $\left(\begin{smallmatrix} 1 & 2 & 3 & 4 & 5 \\ 3 & 5 & 1 & 4 & 2 \end{smallmatrix}\right)$ denotamos a la permutación de $[\![5]\!]$ que manda 1, 2, 3, 4 y 5 a 3, 5, 1, 4 y 2, respectivamente.

4.4.3. Sea $n \in \mathbb{N}$ y sean r y s dos elementos distintos de $[\![n]\!]$. Escribimos $(r\,s)$ a la permutación de $[\![n]\!]$ tal que para cada $i \in [\![n]\!]$ es

$$(r\,s)(i) = \begin{cases} s, & \text{si } i = r; \\ r, & \text{si } i = s; \\ i, & \text{si } i \neq \{r, s\}. \end{cases}$$

Una permutación σ de $[\![n]\!]$ es una ***transposición*** si el conjunto $\{i \in [\![n]\!] : \sigma(i) \neq i\}$ tiene exactamente dos elementos; si ése es el caso y r y s son esos dos elementos, es claro que $\sigma = (r\,s)$.

4.4.4. Proposición. *Sea* $n \in \mathbb{N}$. *Si* σ *es una permutación de* $[\![n]\!]$, *entonces existen* $l \in \mathbb{N}_0$ *y transposiciones* $\tau_1, \dots, \tau_l$ *de* S_n *tales que* $\sigma = \tau_1 \circ \cdots \circ \tau_l$.

Demostración. Para cada $\sigma \in S_n$ sea $|\sigma|$ el conjunto $\{i \in [\![n]\!] : \sigma(i) \neq i\}$ y para cada $k \in \{0, \dots, n\}$ sea $P(k)$ la afirmación

si $\sigma \in S_n$ *y el conjunto* $|\sigma|$ *tiene a lo sumo* k *elementos, entonces existen* $l \in \mathbb{N}_0$ *y transposiciones* $\tau_1, \dots, \tau_l$ *de* S_n *tales que* $\sigma = \tau_1 \circ \cdots \circ \tau_n$. (7)

Como cualquiera sea $\sigma \in S_n$ el conjunto $|\sigma|$ tiene a lo sumo n elementos, para probar a proposición es suficiente con que probemos que la afirmación $P(n)$ vale. Procedemos inductivamente.

Consideremos primero la afirmación $P(0)$. Si σ es una permutación tal que $|\sigma|$ tiene a lo sumo 0 elementos, entonces por supuesto $|\sigma| = \varnothing$ y, en consecuencia, $\sigma(i) = i$ para cada $i \in [\![n]\!]$:

vemos así que $\sigma = \mathrm{id}$ y podemos tomar $l = 0$ en (7), ya que la composición de cero permutaciones es igual a id. Vemos que de esta manera que la afirmación $P(0)$ vale.

Supongamos ahora que $k \in [\![n]\!]$ y que sabemos que $P(k-1)$ vale, y sea σ una permutación de $[\![n]\!]$ tal que $|\sigma|$ tiene a lo sumo k elementos. Si tiene *menos* que k elementos, entonces tiene a lo sumo $k-1$ y, como estamos suponiendo que vale $P(k-1)$, existen $l \in \mathbb{N}_0$ y transposiciones $\tau_1, \dots, \tau_l$ de $[\![n]\!]$ tales que $\sigma = \tau_1 \circ \cdots \circ \tau_l$. Esto muestra que $P(k)$ vale en ese caso.

Nos resta ocuparnos del caso en el que el conjunto $|\sigma|$ tiene exactamente k elementos y, en particular, no es vacío. Sea $i_0 \in |\sigma|$ e $i_1 = \sigma(i_0)$, que es distinto de i_0. Observemos que $i_1 \in |\sigma|$: tendríamos si no que $\sigma(i_1) = i_1$, de manera que $i_1 = \sigma^{-1}(i_1) = i_0$, lo que es absurdo.

Sea $\tau = (i_0\, i_1) \circ \sigma$. Si $i \in [\![n]\!] \setminus |\sigma|$, entonces

$$\tau(i) = (i_0\, i_1)\big(\sigma(i)\big) = (i_0\, i_1)(i) = i,$$

ya que $i \notin \{i_0, i_1\}$. Esto significa que $|\tau| \subseteq |\sigma|$. Por otro lado, es

$$\tau(i_0) = (i_0\, i_1)\big(\sigma(i_0)\big) = (i_0\, i_1)(i_1) = i_0,$$

así que $i_0 \notin |\tau|$. Vemos de esta forma que $|\tau|$ está estrictamente contenido en $|\sigma|$ y que, en particular, tiene a lo sumo $k-1$ elementos. Como estamos suponiendo que vale la afirmación $P(k-1)$, esto implica que existen $l \in \mathbb{N}$ y transposiciones $\tau_1, \dots, \tau_l$ de $[\![n]\!]$ tales que $\tau = \tau_1 \circ \cdots \circ \tau_l$. Como $(i_0\, i_1) \circ (i_0\, i_1) = \mathrm{id}$, se tiene entonces que

$$\sigma = (i_0\, i_1) \circ (i_0\, i_1) \circ \sigma = (i_0\, i_1) \circ \tau = (i_0\, i_1) \circ \tau_1 \circ \cdots \circ \tau_l,$$

esto es, que σ es igual a una composición de transposiciones. Esto prueba que la afirmación $P(k)$ también vale en este caso, y completa la demostración de la proposición. □

4.4.5. La Proposición 4.4.4 nos dice que toda permutación puede escribirse como una composición de transposiciones. Es importante observar que esta escritura no es única: ni las transposiciones que aparecen en ella ni su cantidad están determinadas. Es fácil verificar, por ejemplo, que en S_4 se tiene que

$$(1\,3) = (1\,2) \circ (2\,3) \circ (1\,2) = (2\,3) \circ (1\,2) \circ (2\,3) = (2\,3) \circ (1\,4) \circ (2\,4) \circ (2\,3) \circ (1\,4).$$

4.4.6. Si $n \in \mathbb{N}$ y $\sigma \in S_n$ es una permutación, la ***matriz de permutación*** asociada a σ es la matriz $A(\sigma) = (a_{i,j}) \in \mathsf{M}_n(\Bbbk)$ tal que para cada $i, j \in [\![n]\!]$ es

$$a_{i,j} = \begin{cases} 1, & \text{si } \sigma(j) = i; \\ 0, & \text{en caso contrario.} \end{cases}$$

Por ejemplo, la matriz de permutación correspondiente a la permutación 3 5 1 4 2 de $[\![5]\!]$ es

$$\begin{pmatrix} 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 & 0 \end{pmatrix}$$

4.4.7. Proposición. *Sea $n \in \mathbb{N}$ y sea σ una permutación de $[\![n]\!]$. Para cada $i \in [\![n]\!]$ se tiene que*

$$A(\sigma)e_i = e_{\sigma(i)}.$$

Demostración. Sea $i \in [\![n]\!]$ y sea $A(\sigma) = (a_{i,j})$. Si $k \in [\![n]\!]$, la componente k-ésima del vector $A(\sigma)e_i$ es $a_{i,k}$, y esto es 1 si $k = \sigma(i)$ y 0 en caso contrario. Esto nos dice que $A(\sigma)e_i$ es el vector $e_{\sigma(i)}$, como afirma la proposición. □

4.4.8. Proposición. *Sea $n \in \mathbb{N}$.*

(*i*) *La matriz de permutación $A(\mathrm{id})$ correspondiente a la permutación identidad* id *de $[\![n]\!]$ es la matriz identidad de $\mathsf{M}_n(\mathbb{k})$.*

(*ii*) *Si σ y τ son permutaciones de $[\![n]\!]$, entonces*

$$A(\sigma \circ \tau) = A(\sigma) \cdot A(\tau).$$

(*iii*) *Si σ es una permutación de $[\![n]\!]$, entonces la matriz $A(\sigma)$ es inversible y*

$$A(\sigma)^{-1} = A(\sigma^{-1}) = A(\sigma)^{\mathrm{t}}.$$

Demostración. (*i*) Sea $A = (a_{i,j})$ la matriz de permutación de $\mathrm{id} \in S_n$. De acuerdo a la definición, para cada $i, j \in [\![n]\!]$ el escalar $a_{i,j}$ es igual a 1 si $j = i$ y a 0 en caso contrario: esto significa precisamente que A es a matriz identidad de $\mathsf{M}_n(\mathbb{k})$.

(*ii*) Sean σ y τ permutaciones de $[\![n]\!]$ y sean $A(\sigma) = (a_{i,j})$ y $A(\tau) = (b_{i,j})$ las correspondientes matrices de permutación. Sea $(c_{i,j})$ la matriz producto $A(\sigma) \cdot A(\tau)$. Si $i, j \in [\![n]\!]$, entonces

$$c_{i,j} = \sum_{k=1}^{n} a_{i,k} b_{k,j}.$$

El escalar $b_{k,j}$ es nulo salvo si $k = \tau(j)$, y en ese caso es igual a 1: vemos así que, de hecho, $c_{i,j} = a_{i,\tau(j)}$, y esto nos dice que $c_{i,j} = 1$ si $i = \sigma(\tau(j))$ y que $c_{i,j} = 0$ en caso contrario. Por supuesto, esto significa que la matriz $(c_{i,j})$ es precisamente la matriz de permutación correspondiente a la permutación $\sigma \circ \tau$, y esto es lo que se afirma en la proposición.

(*iii*) Sea σ una permutación de $[\![n]\!]$. De acuerdo a las partes (*ii*) y (*i*) de la proposición, que ya probamos, tenemos que

$$A(\sigma^{-1}) \cdot A(\sigma) = A(\sigma^{-1} \circ \sigma) = A(\mathrm{id}) = I_n,$$

la matriz identidad de $\mathsf{M}_n(\mathbb{k})$ y, de manera similar, que $A(\sigma) \cdot A(\sigma^{-1}) = I_n$. Así, $A(\sigma)$ y $A(\sigma^{-1})$ son matrices inversas y, en particular, la matriz $A(\sigma)$ es inversible.

Por otro lado, supongamos que $A(\sigma) = (a_{i,j})$ y sea $(d_{i,j}) = A(\sigma) \cdot A(\sigma)^{\mathrm{t}}$. Si $i, j \in [\![n]\!]$, entonces

$$d_{i,j} = \sum_{k=1}^{n} a_{i,k} a_{j,k}$$

Ahora bien: para cada $k \in [\![n]\!]$ es escalar $a_{k,j}$ es nulo salvo si $j = \sigma(k)$ y $a_{i,k}$ es nulo salvo si $i = \sigma(k)$ y, recordando que σ es una función biyectiva, se sigue de esto que hay dos casos: o bien $i = j$ y en ese caso $d_{i,j} = 1$, o bien $i \neq j$ y en ese caso $d_{i,j} = 0$. Vemos de esta forma que $A(\sigma) \cdot A(\sigma)^{\mathrm{t}} = I_n$ y un razonamiento similar muestra que $A(\sigma)^{\mathrm{t}} \cdot A(\sigma) = I_n$: podemos concluir entonces, como queremos, que $A(\sigma)^{\mathrm{t}} = A(\sigma)^{-1}$. □

4.4.9. Llamamos ***signo*** de una permutación $\sigma \in S_n$, y escribimos $\operatorname{sgn}(\sigma)$, al valor del determinante de la matriz de permutación asociada a σ, esto es,

$$\operatorname{sgn}(\sigma) = \det A(\sigma).$$

4.4.10. Proposición. *Sea $n \in \mathbb{N}$.*

(*i*) *El signo de la permutación identidad es* $\operatorname{sgn}(\mathrm{id}) = 1$.

(*ii*) *Si σ y τ son permutaciones de $[\![n]\!]$, entonces*

$$\operatorname{sgn}(\sigma \circ \tau) = \operatorname{sgn}(\sigma) \cdot \operatorname{sgn}(\tau).$$

Demostración. Como la matriz de permutación correspondiente a la permutación identidad id de S_n es la matriz identidad I_n, es $\operatorname{sgn}(\mathrm{id}) = \det I_n = 1$. Por otro lado, si σ y τ son elementos de S_n, entonces

$$\operatorname{sgn}(\sigma \circ \tau) = \det A(\sigma \circ \tau) = \det A(\sigma) \cdot A(\tau) = \det A(\sigma) \cdot \det A(\tau) = \operatorname{sgn}(\sigma) \cdot \operatorname{sgn}(\tau),$$

por la Proposición 4.4.8*(ii)* y la Proposición 4.3.3*(ii)*. □

4.4.11. Proposición. *Sea $n \in \mathbb{N}$ y sea σ una permutación de $[\![n]\!]$. Si $l \in \mathbb{N}_0$ y $\tau_1, \dots, \tau_l$ son transposiciones de $[\![n]\!]$ tales que $\sigma = \tau_1 \circ \cdots \circ \tau_l$, entonces* $\operatorname{sgn}(\sigma) = (-1)^l$.

Demostración. Sea σ una permutación de $[\![n]\!]$ y sean $l \in \mathbb{N}$ y $\tau_1, \dots, \tau_l$ son transposiciones de $[\![n]\!]$ tales que $\sigma = \tau_1 \circ \cdots \circ \tau_l$. De acuerdo a la Proposición 4.4.10*(ii)*, tenemos que

$$\operatorname{sgn}(\sigma) = \operatorname{sgn}(\tau_1) \cdots \operatorname{sgn}(\tau_l).$$

Para probar la proposición, entonces, bastará que probemos que

> *si τ es una transposición de $[\![n]\!]$, entonces* $\operatorname{sgn}(\tau) = -1$.

Sean entonces r y s dos elementos de $[\![n]\!]$ tales que $r < s$ y sea $\tau = (r\,s)$ la transposición que los intercambia. Para cada $i \in [\![n]\!]$ es

$$A(\tau)e_i = e_{\tau(i)} = \begin{cases} e_i & \text{si } i \notin \{r, s\}; \\ e_s & \text{si } i = r; \\ e_s & \text{si } i = s. \end{cases}$$

Nuestras definiciones nos dicen entonces que

$$\begin{aligned} \operatorname{sgn}(\tau) &= \det A(\sigma) = D_n(A(\sigma)e_1, \dots, A(\sigma)e_n) = D_n(e_1, \dots, \underbrace{e_s}_{r}, \dots, \underbrace{e_r}_{s}, \dots, e_n) \\ &= -D_n(e_1, \dots, e_n) = -1, \end{aligned}$$

como queremos. □

4.4.12. Una primera aplicación de la Proposición 4.4.11 el siguiente corolario:

Corolario. *Sea $n \in \mathbb{N}$. Si σ es una permutación de $[\![n]\!]$, entonces* $\operatorname{sgn}(\sigma) \in \{+1, -1\}$ *y*

$$\operatorname{sgn}(\sigma^{-1}) = \operatorname{sgn}(\sigma).$$

Demostración. La primera afirmación sigue inmediatamente de la Proposición 4.4.11. Por otro lado, sabemos de la Proposición 4.4.10 que

$$\operatorname{sgn}(\sigma) \cdot \operatorname{sgn}(\sigma^{-1}) = \operatorname{sgn}(\sigma \circ \sigma^{-1}) = \operatorname{sgn}(\mathrm{id}) = 1$$

y esto, junto a la primera afirmación, implica la segunda. □

4.4.13. Un segundo corolario de la Proposición 4.4.11 es el siguiente:

Corolario. *Sea $n \in \mathbb{N}$. Si k y l son enteros no negativos y $\tau_1, \dots, \tau_k$ y $\rho_1, \dots, \rho_l$ son transposiciones de S_n tales que $\tau_1 \circ \cdots \circ \tau_k = \rho_1 \circ \cdots \circ \rho_l$, entonces los números k y l tienen la misma paridad.*

Decimos que una permutación σ de $[\![n]\!]$ es ***par*** si es igual a la composición de un número par de transposiciones y que es ***impar*** en caso contrario: el corolario nos dice que esta definición tiene sentido, ya que la paridad del número de factores en una escritura de σ como producto de transposiciones depende solamente de σ y no de la escritura elegida. Así, la permutación 3 5 1 4 2 de $[\![5]\!]$ es par, porque es igual a $(1\,3) \circ (2\,5)$, mientras que la permutación 2 4 1 3 5 es impar, porque es igual a $(1\,2) \circ (2\,4) \circ (3\,4)$.

Demostración. Si k, l, $\tau_1, \dots, \tau_k$ y $\rho_1, \dots, \rho_l$ son como en el enunciado, la Proposición 4.4.11 nos dice que $(-1)^k = (-1)^l$ y el corolario sigue inmediatamente de esto. □

4.4.14. Estamos por fin en posición dar una fórmula explícita para el determinante de una matriz:

Proposición. *Sea $n \in \mathbb{N}$. Si $A = (a_{i,j}) \in \mathsf{M}_n(\Bbbk)$, entonces*

$$\det A = \sum_{\sigma \in S_n} \operatorname{sgn}(\sigma)\, a_{\sigma(1),1} \cdots a_{\sigma(n),n}.$$

Esta expresión para el determinante de una matriz se llama la ***fórmula de Leibniz***, por Gottfried Wilhelm Leibniz (1646–1716, Alemania). Observemos que la suma tiene $n!$ términos.

Demostración. Sea $A = (a_{i,j})$ un elemento de $\mathsf{M}_n(\Bbbk)$ y sea $(e_1, \dots, e_n)$ la base ordenada estándar de $\Bbbk^n$. Para cada $i \in [\![n]\!]$ es $Ae_i = a_{1,i}e_1 + \cdots + a_{n,i}e_n$, así que

$$\det A = D_n(a_{1,1}e_1 + \cdots + a_{n,1}e_n, \dots, a_{1,n}e_1 + \cdots + a_{n,n}e_n)$$

y, como la función D_n es multilineal, esto es igual a

$$= \sum_{i_1,\dots,i_n \in [\![n]\!]} a_{i_1,1} \cdots a_{i_n,n} D_n(e_{i_1}, \dots, e_{i_n}), \tag{8}$$

con la suma tomada sobre todas las formas de elegir los n índices $i_1, \dots, i_n$ en el conjunto $[\![n]\!]$.

Ahora bien: si $i_1, \dots, i_n$ son elementos de $[\![n]\!]$ que no distintos dos a dos, el hecho de que la función D_n es alternante implica que $D_n(e_{i_1}, \dots, e_{i_n}) = 0$. Esto significa que en la suma (8) basta considerar sólo los términos que corresponden a elecciones de índices $i_1, \dots, i_n$ en $[\![n]\!]$ que distintos dos a dos, esto es, tales que $\left(\begin{smallmatrix} 1 & \cdots & n \\ i_1 & \cdots & i_n \end{smallmatrix}\right)$ sea una permutación de $[\![n]\!]$, y entonces vemos que

$$\det A = \sum_{\sigma \in S_n} a_{\sigma(1),1} \cdots a_{\sigma(n),n} D_n(e_{\sigma(1)}, \dots, e_{\sigma(n)}). \tag{9}$$

Para cada $\sigma \in S_n$ sabemos que $e_{\sigma(j)} = A(\sigma)e_j$ si $j \in [\![n]\!]$, así que

$$D_n(e_{\sigma(1)}, \dots, e_{\sigma(n)}) = D_n(A(\sigma)e_1, \dots, A(\sigma)e_n) = \det A(\sigma) = \operatorname{sgn}(\sigma).$$

Usando esto en el lado derecho de la igualdad (9) obtenemos la fórmula que aparece en el enunciado de la proposición. □

4.4.15. Un corolario casi inmediato de la fórmula de Leibniz es el siguiente:

Corolario. *Sea $n \in \mathbb{N}$*

(*i*) *Si $A \in \mathsf{M}_n(\mathbb{Q})$ y todas las entradas de A están en $\mathbb{Z}$, entonces* $\det A \in \mathbb{Z}$.

(*ii*) *Si $A \in \mathsf{M}_n(\mathbb{k}(X))$ y todas las entradas de A están en $\mathbb{k}[X]$, entonces* $\det A \in \mathbb{k}[X]$.

Demostración. **Hacer.** □

§5. El determinante de la matriz transpuesta de una matriz

4.5.1. Usando la fórmula de Leibniz podemos obtener fácilmente el siguiente resultado, que es importante tanto teóricamente como en la práctica: una matriz tiene el mismo determinante que su transpuesta.

Proposición. *Sea $n \in \mathbb{N}$. Para cada $A \in \mathsf{M}_n(\mathbb{k})$ se tiene que* $\det A = \det A^{\mathsf{t}}$.

Demostración. Sea $A = (a_{i,j}) \in \mathsf{M}_n(\mathbb{k})$. De acuerdo a la Proposición 4.4.14, es

$$\det A = \sum_{\sigma \in S_n} \operatorname{sgn}(\sigma)\, a_{\sigma(1),1} \cdots a_{\sigma(n),n}.$$

La función $\sigma \in S_n \mapsto \sigma^{-1} \in S_n$ es una biyección —es su propia función inversa, de hecho— así que podemos reescribir esta suma en la forma

$$\sum_{\tau \in S_n} \operatorname{sgn}(\tau^{-1})\, a_{\tau^{-1}(1),1} \cdots a_{\tau^{-1}(n),n}. \tag{10}$$

Si τ es una permutación de $[\![n]\!]$ tenemos que $\operatorname{sgn}(\tau^{-1}) = \operatorname{sgn}(\tau)$ y el producto

$$a_{\tau^{-1}(1),1} \cdots a_{\tau^{-1}(n),n}$$

coincide con

$$a_{1,\tau(1)} \cdots a_{n,\tau(n)},$$

ya que tiene exactamente los mismos n factores aunque en otro orden. La suma (10) tiene el mismo valor, entonces, que

$$\sum_{\tau\in S_n} \operatorname{sgn}(\tau) a_{1,\tau(1)} \cdots a_{n,\tau(n)}.$$

Si $(b_{i,j})$ es la matriz transpuesta A^{t}, de manera que $b_{i,j} = a_{j,i}$ para cada elección de i, $j \in [\![n]\!]$, podemos escribir esto en la forma

$$\sum_{\tau\in S_n} \operatorname{sgn}(\tau) b_{\tau(1),1} \cdots b_{\tau(n),n}$$

y es claro, ahora, que esto es igual a $\det A^{\mathrm{t}}$. Esto prueba la proposición. □

4.5.2. Una consecuencia de la Proposición 4.5.1 es que si en el enunciado de la Proposición 4.3.2 reemplazamos la palabra *columna* por *fila* obtenemos también una afirmación cierta:

Proposición. *Sea $n \in \mathbb{N}$. La función $\det : \mathsf{M}_n(\Bbbk) \to \Bbbk$ tiene las siguientes propiedades:*

(i) $\det A$ *depende multilinealmente de las filas de la matriz A.*

(ii) *Si $A \in \mathsf{M}_n(\Bbbk)$ tiene dos filas iguales, entonces* $\det A = 0$.

(iii) *Si A y B son elementos de $\mathsf{M}_n(\Bbbk)$ y B se obtiene de A intercambiando dos filas, entonces* $\det B = -\det A$.

(iv) *Si A y B son elementos de $\mathsf{M}_n(\Bbbk)$ y B se obtiene de A sumando a una de sus filas una combinación lineal de las demás, entonces* $\det B = \det A$.

Demostración. Estas afirmaciones se deducen de las correspondientes en la Proposición 4.3.2 usando el hecho de que el determinante de una matriz coincide con el de su matriz transpuesta, y la observación evidente de que las filas de aquélla son las filas de ésta, a menos de transponer. □

§6. La fórmula de Laplace

4.6.1. Si m, $n \in \mathbb{N}$ y $A \in \mathsf{M}_{m,n}(\Bbbk)$ es una matriz, llamamos ***menor*** de A a toda matriz que se obtiene de A eliminando algún número de filas y de columnas. Cuando m, $n \geq 2$, $r \in [\![m]\!]$ y $s \in [\![n]\!]$, escribimos $A^{(r,s)}$ al menor de A que se obtiene eliminando la fila r-ésima y la columna s-ésima, que es un elemento de $\mathsf{M}_{m-1,n-1}(\Bbbk)$.

4.6.2. Proposición. *Sea $n \in \mathbb{N}$ y sea $A = (a_{i,j})$ un elemento de $\mathsf{M}_(\Bbbk)$.*

(i) *Para cada $i \in [\![n]\!]$ se tiene que*

$$\det A = \sum_{l=1}^{n} (-1)^{i+l} a_{i,l} \cdot \det A^{(i,l)}.$$

(*ii*) *Para cada* $j \in [\![n]\!]$ *se tiene que*

$$\det A = \sum_{l=1}^{n} (-1)^{j+l} a_{l,j} \cdot \det A^{(l,j)}.$$

Estas fórmulas para el determinante de una matriz se llaman *fórmulas de Laplace* o *desarrollos de* $\det A$ *a lo largo de la fila* i*-ésima de* A y *de la columna* j*-esima*, respectivamente. Si $i, j \in [\![n]\!]$, llamamos al escalar $(-1)^{i+j} \det A^{(i,j)}$ el ***cofactor*** de la entrada $a_{i,j}$ de A.

Demostración. Basta que probemos la primera parte de la proposición, ya que la segunda se obtiene inmediatamente de la primera considerando la matriz transpuesta de A.

Ahora bien, que vale (*i*) cuando i es 1 es consecuencia de la forma en que definimos la función D_n en la prueba de la Proposición 4.2.4. Si i no es 1, entonces podemos considerar la matriz B que se obtiene de A intercambiando su primera fila y la i-ésima. De la Proposición 4.5.2*(iii)* sabemos que $\det B = -\det A$. Por otro lado, para cada $l \in [\![n]\!]$ el menor $B^{(1,l)}$ es la matriz que se obtiene del menor $A^{(i,l)}$ intercambiando su primera fila con la segunda, luego la segunda con la tercera y así hasta intercambiar la $(i-2)$-ésima por la $(i-1)$-ésima: estos son $i-2$ intercambios, así que $\det B^{(1,l)} = (-1)^i \det A^{(i,l)}$. Teniendo todo esto en cuenta, concluimos que

$$\begin{aligned}\det A = -\det B &= -\sum_{l=1}^{n} (-1)^{l+1} b_{1,l} \cdot \det B^{(1,l)} = -\sum_{l=1}^{n} (-1)^{l+i+1} a_{i,l} \cdot \det A^{(1,l)} \\ &= \sum_{l=1}^{n} (-1)^{l+i} a_{i,l} \cdot \det A^{(1,l)},\end{aligned}$$

como queremos. □

4.6.3. Como consecuencia de las fórmulas de Laplace y la alternancia del determinante, obtenemos el siguiente corolario:

Corolario. *Sea* $n \in \mathbb{N}$ *y sea* $A = (a_{i,j})$ *un elemento de* $\mathsf{M}_n(\Bbbk)$.

(*i*) *Si* $i, i' \in [\![n]\!]$ *son distintos, entonces*

$$\sum_{l=1}^{n} (-1)^{i+l} a_{i,l} \cdot \det A^{(i',l)} = 0.$$

(*ii*) *Si* $j, j' \in [\![n]\!]$ *son distintos, entonces*

$$\sum_{l=1}^{n} (-1)^{j+l} a_{l,j} \cdot \det A^{(l,j')} = 0.$$

Demostración. Otra vez, gracias a la Proposición 4.5.1, es suficiente que probemos la primera parte. Sean i e i' elementos distintos de $[\![n]\!]$ y sea $B = (b_{i,j})$ la matriz que se obtiene de la matriz A reemplazando su fila i'-ésima por una copia de su fila i-ésima. Como B tiene dos filas iguales, sabemos que $\det B = 0$ y entonces el desarrollo de Laplace para el determinante de B a lo largo de la fila i'-ésima de esta matriz nos dice que

$$\sum_{l=1}^{n} (-1)^{i+l} b_{i',l} \cdot \det B^{(i',l)} = 0. \tag{11}$$

Si $l \in [\![n]\!]$, es $b_{i',l} = a_{i,l}$ y el menor $B^{(i',l)}$ coincide con el menor $A^{(i',l)}$: la igualdad (11) implica entonces que

$$\sum_{l=1}^{n} (-1)^{i+l} a_{i,l} \cdot \det A^{(i',l)} = 0,$$

como afirma la proposición. □

§7. La regla de Cramer

4.7.1. La Proposición 4.3.3*(iii)* nos dice que una condición necesaria para que una matriz sea inversible es que su determinante sea no nulo. Podemos probar ahora que esta condición es también suficiente:

Proposición. *Sea $n \in \mathbb{N}$, sea $A = (a_{i,j}) \in \mathsf{M}_n(\mathbb{k})$ y consideremos la matriz* $\operatorname{adj} A = (b_{i,j})$ *tal que* $b_{i,j} = (-1)^{i+j} \det A^{(j,i)}$ *para cada i, $j \in [\![n]\!]$.*

(*i*) *Se tiene que*

$$A \cdot \operatorname{adj} A = \operatorname{adj} A \cdot A = \det A \cdot I_n.$$

(*ii*) *La matriz A es inversible si y solamente si su determinante* $\det A$ *es no nulo, y en ese caso vale que*

$$A^{-1} = \frac{1}{\det A} \operatorname{adj} A. \tag{12}$$

Demostración. Si i, $j \in [\![n]\!]$, la entrada (i, j)-ésima del producto $A \cdot \operatorname{adj} A$ es

$$\sum_{l=1}^{n} a_{i,l} b_{l,j} = \sum_{l=1}^{n} a_{i,l} (-1)^{l+j} \det A^{(j,l)}$$

y de acuerdo a las Proposiciones 4.6.2*(i)* y 4.6.3*(i)*, esto es igual a $\det A$ si $i = j$ y a 0 si $i \neq j$. Esto significa que $A \cdot \operatorname{adj} A = \det A \cdot I_n$. De la misma forma, a partir de las Proposiciones 4.6.2*(ii)* y 4.6.3*(ii)* podemos ver que $\operatorname{adj} A \cdot A = \det A \cdot I_n$. Esto prueba la afirmación (*i*) de la proposición.

Ya sabemos de la Proposición 4.3.3*(iii)* que es necesario para que la matriz A sea inversible que su determinante sea no nulo. Recíprocamente, si tenemos que $\det A \neq 0$, entonces en la igualdad de (*i*) que acabamos de probar podemos dividir por el escalar $\det A$ para ver que

$$A \cdot \frac{\operatorname{adj} A}{\det A} = \frac{\operatorname{adj} A}{\det A} \cdot A = I_n$$

y concluir que A es inversible y que su matriz inversa es la descripta en (12). □

4.7.2. Usando la expresión para la inversa de una matriz inversible que nos da la Proposición 4.7.1 podemos obtener una expresión cerrada para solución de un sistema de ecuaciones con la misma cantidad de incógnitas que de ecuaciones con matriz de coeficientes inversible. Llamamos a este resultado la *regla de Cramer*, por Daniel Cramer (1704–1752, Suiza).

Proposición. *Sean $n \in \mathbb{N}$, $A = (a_{i,j}) \in \mathsf{M}_n(\Bbbk)$ y $b = (b_i) \in \Bbbk^n$. Si la matriz A es inversible, de manera que* $\det A \neq 0$*, entonces existe exactamente un vector $x = (x_i) \in \Bbbk^n$ tal que $Ax = b$ y para cada $i \in [\![n]\!]$ la componente i-ésima de x es*

$$x_i = \frac{\begin{vmatrix} a_{1,1} & \cdots & a_{1,i-1} & b_1 & a_{1,i+1} & \cdots & a_{1,n} \\ \vdots & & \vdots & \vdots & \vdots & & \vdots \\ a_{n,1} & \cdots & a_{n,i-1} & b_n & a_{n,i+1} & \cdots & a_{n,n} \end{vmatrix}}{\det A}$$

El numerador de este cociente es el determinante de la matriz que se obtiene de A reemplazando su columna i-ésima por el vector b.

Demostración. Supongamos que la matriz A es inversible. La función lineal $f : x \in \Bbbk^n \mapsto Ax \in \Bbbk^n$ es entonces un isomorfismo y su inversa es la función $g : x \in \Bbbk^n \mapsto A^{-1}x \in \Bbbk^n$. En particular, como f es biyectiva, existe exactamente un elemento x de $\Bbbk^n$ tal que $Ax = f(x) = b$ y es $x = g(x) = A^{-1}x$. En vista de la Proposición 4.7.1*(ii)*, esto significa que

$$x = \frac{1}{\det A} \operatorname{adj} A \cdot x.$$

Si $i \in [\![n]\!]$ y escribimos $\operatorname{adj} A = (b_{i,j})$, la componente i-ésima de este vector es

$$\frac{1}{\det A} \sum_{l=1}^{n} b_{i,l} x_l = \sum_{l=1}^{n} (-1)^{i+l} x_l \det A^{(l,i)}$$

y la suma que aparece aquí es el desarrollo de Laplace a lo largo de la columna i-ésima del determinante de la matriz que se obtiene de A reemplazando la columna i-ésima por el vector b. Esto es precisamente lo que afirma la proposición. □

§8. El rango de una matriz

4.8.1. El determinante nos da un criterio muy compacto para decidir la independencia lineal de n vectores de $\Bbbk^n$:

Proposición. *Sea $n \in \mathbb{N}$ y sean $x_1, \dots, x_n \in \Bbbk^n$. Los vectores $x_1, \dots, x_n$ son linealmente independientes si y solamente si la matriz de $\mathsf{M}_n(\Bbbk)$ que los tiene por columnas tiene determinante no nulo.*

Demostración. Sea $A \in \mathsf{M}_n(\Bbbk)$ la matriz que tiene a los vectores $x_1, \dots, x_n$ por columnas. Sabemos que si los vectores son linealmente dependientes, entonces el determinante de A es nulo. Recíprocamente, si son linealmente independientes, entonces la función $f : x \in \Bbbk^n \mapsto Ax \in \Bbbk^n$ es inversible y, de acuerdo a la Proposición 2.8.6, la matriz $[f]^{\mathscr{B}}_{\mathscr{B}}$ de f con respecto a la base estándar $\mathscr{B}$ de $\Bbbk^n$, que coincide con A, como vimos en el Ejemplo 2.8.2(a), es inversible. En vista de la Proposición 4.7.1*(ii)*, entonces, es $\det A \neq 0$. □

4.8.2. Usando este resultado podemos obtener una nueva descripción del rango de una matriz, esta vez en términos del tamaño de sus menores cuadrados no singulares:

Proposición. *Sean m, $n \in \mathbb{N}$ y sea $A \in \mathsf{M}_{m,n}(\Bbbk)$. El rango de A es igual al tamaño máximo de los menores cuadrados de A que tienen determinante no nulo.*

Demostración. Sea r el rango de la matriz A y sea s el tamaño máximo de un menor cuadrado de A con determinante no nulo. Como r es la dimensión del subespacio de $\Bbbk^m$ generado por las columnas de A, sabemos que hay r columnas en A que son linealmente independientes: sea B el menor de A que se obtiene eliminando todas las otras $n - r$ columnas, de manera que $B \in \mathsf{M}_{m,r}(\Bbbk)$. Por construcción, las columnas de B son linealmente independientes, así que el rango de B es r. De acuerdo a la Proposición 3.4.7, el rango por filas de B también es r y, en consecuencia, hay r filas en B que son linealmente independientes. Sea C el menor de B que se obtiene eliminando las otras $m - r$ filas. Es claro que C es también un menor de la matriz A con la que empezamos, se trata de un menor cuadrado de tamaño r y, como sus filas son linealmente independientes, su determinante es no nulo: la elección de s, en consecuencia, implica que es $s \geq r$.

Por otro lado, si D es un menor cuadrado de A de tamaño *mayor* que r, entonces las filas de A que tocan a D son linealmente dependientes —ya que A no posee $r + 1$ columnas linealmente independientes— y entonces las columnas de D son también linealmente dependientes: esto implica que $\det D = 0$. Así, todo menor cuadrado de A de tamaño mayor que r tiene determinante nulo y, por lo tanto, $s \leq r$. □

§9. Tres determinantes

Matrices triangulares

4.9.1. Decimos que una matriz $A = (a_{i,j}) \in \mathsf{M}_n(\Bbbk)$ es ***triangular superior*** si cada vez que $i, j \in [\![n]\!]$ son tales que $i > j$ se tiene que $a_{i,j} = 0$, esto es, si todas las entradas de la matriz que están por debajo de la diagonal principal son nulas. De forma simétrica, decimos que A es ***triangular inferior*** si cada vez que $i, j \in [\![n]\!]$ son tales que $i < j$ se tiene que $a_{i,j} = 0$

Proposición. *Sea $n \in \mathbb{N}$. Si $A = (a_{i,j}) \in \mathsf{M}_n(\Bbbk)$ es una matriz triangular superior o triangular inferior, entonces el determinante de A es el producto de las entradas que aparecen en su diagonal,* *esto es,*

$$\det A = a_{1,1} \cdots a_{n,n}.$$

Demostración. Basta que consideremos el caso en que A es triangular superior, ya que si A es triangular inferior entonces la matriz transpuesta A^{t} es triangular superior y, como sabemos, $\det A = \det A^{\mathsf{t}}$.

Supongamos entonces que la matriz A es triangular superior. Según la Proposición 4.4.14, es

$$\det A = \sum_{\sigma \in S_n} \operatorname{sgn}(\sigma) a_{\sigma(1),1} \cdots a_{\sigma(n),n}. \tag{13}$$

Si σ es una permutación de $[\![n]\!]$, entonces el producto $a_{\sigma(1),1} \cdots a_{\sigma(n),n}$ es nulo si cualquiera de sus n factores es nulo, y esto ocurre si existe $i \in [\![n]\!]$ tal que $\sigma(i) > i$. Vemos así que la suma (13) no cambia si solamente sumamos los términos que corresponden a permutaciones $\sigma \in S_n$ tales que $\sigma(i) \leq i$ para cada $i \in [\![n]\!]$. Ahora bien: hay exactamente una permutación que satisface esta condición, la permutación identidad, y, en consecuencia, tenemos que

$$\det A = \operatorname{sgn}(\mathrm{id}) a_{\mathrm{id}(1),1} \cdots a_{\mathrm{id}(n),n} = a_{1,1} \cdots a_{n,n},$$

como afirma la proposición. □

4.9.2. Más generalmente, podemos describir el determinante de una matriz triangular por bloques:

Proposición. *Sean n, $r \in \mathbb{N}$, sean $n_1, \dots, n_r \in \mathbb{N}$ tales que $n = n_1 + \cdots + n_r$ y supongamos que para cada i, $j \in [\![r]\!]$ con $i \leq j$ tenemos una matriz $A_{i,j} \in \mathsf{M}_{n_i,n_j}(\Bbbk)$. Si A es la matriz de bloques*

$$\begin{pmatrix} A_{1,1} & A_{1,2} & \cdots & A_{1,r-1} & A_{1,r} \\ 0 & A_{2,2} & \cdots & A_{2,r-1} & A_{2,r} \\ \vdots & & \ddots & \vdots & \vdots \\ \vdots & & & A_{r-1,r-1} & A_{r-1,r} \\ 0 & \cdots & 0 & 0 & A_{r,r} \end{pmatrix},$$

entonces el determinante de A es

$$\det A = \det A_{1,1} \cdots \det A_{r,r}.$$

Demostración. Como podemos hacer una inducción evidente con respecto a r, es suficiente que consideremos el caso en que $r = 2$. Sean entonces n, n_1, $n_2 \in \mathbb{N}$ tales que $n = n_1 + n_2$, sean $A_{1,1} \in \mathsf{M}_{n_1,n_1}(\Bbbk)$, $A_{1,2} \in \mathsf{M}_{n_1,n_2}(\Bbbk)$ y $A_{2,2} \in \mathsf{M}_{n_2,n_2}(\Bbbk)$, consideremos la matriz de bloques

$$A = \begin{pmatrix} A_{1,1} & A_{1,2} \\ 0 & A_{2,2} \end{pmatrix}.$$

y mostremos que

$$\det A = \det A_{1,1} \cdot \det A_{2,2}. \tag{14}$$

Si la matriz $A_{1,1}$ tiene determinante nulo, sus filas son linealmente dependientes y entonces claramente las primeras n_1 columnas de la matriz A son linealmente dependientes: vemos que en este caso $\det A = 0$ y que, en consecuencia, vale la igualdad (14).

Supongamos ahora que $A_{1,1}$ tiende determinante no nulo, de manera que es inversible. Vale entonces que

$$A = \begin{pmatrix} A_{1,1} & 0 \\ 0 & I_{n_2} \end{pmatrix} \cdot \begin{pmatrix} I_{n_1} & A_{1,1}^{-1}A_{1,2} \\ 0 & A_{2,2} \end{pmatrix}$$

y, en consecuencia, que

$$\det A = \det \begin{pmatrix} A_{1,1} & 0 \\ 0 & I_{n_2} \end{pmatrix} \cdot \det \begin{pmatrix} I_{n_1} & A_{1,1}^{-1}A_{1,2} \\ 0 & A_{2,2} \end{pmatrix}. \tag{15}$$

Ahora bien, desarrollando el determinante por la última fila vemos inmediatamente que

$$\det \begin{pmatrix} A_{1,1} & 0 \\ 0 & I_{n_2} \end{pmatrix} = \det \begin{pmatrix} A_{1,1} & 0 \\ 0 & I_{n_2-1} \end{pmatrix}$$

y entonces una inducción evidente muestra que, de hecho,

$$\det \begin{pmatrix} A_{1,1} & 0 \\ 0 & I_{n_2} \end{pmatrix} = \det A_{1,1}.$$

De manera similar, desarrollando el determinante por la primera columna y haciendo inducción en n_1 vemos que

$$\det \begin{pmatrix} I_{n_1} & A_{1,1}^{-1}A_{1,2} \\ 0 & A_{2,2} \end{pmatrix} = \det A_{2,2}$$

Estas últimas dos igualdades y (15) implican que vale la igualdad (14), como queremos. □

Matrices compañeras

4.9.3. Si $n \in \mathbb{N}$ y $p(X) = c_0 + c_1X + \cdots + c_{n-1}X^{-1} + X^n \in \Bbbk[X]$ es un polinomio mónico de grado n, la ***matriz compañera*** de p es la matriz

$$C(p) = \begin{pmatrix} 0 & 0 & \dots & 0 & -c_0 \\ 1 & 0 & \dots & 0 & -c_1 \\ 0 & 1 & \dots & 0 & -c_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \dots & 1 & -c_{n-1} \end{pmatrix} \in \mathsf{M}_n(\Bbbk),$$

de manera que si $C(p) = (a_{i,j})$, entonces para cada $i, j \in [\![n]\!]$ es

$$a_{i,j} = \begin{cases} 1, & \text{si } i = j+1; \\ -c_{i-1}, & \text{si } j = n; \\ 0, & \text{en cualquier otro caso.} \end{cases}$$

Proposición. *Sea $n \in \mathbb{N}$ y sea $p(X) = c_0 + c_1 X + \cdots + c_{n-1}X^{-1} + X^n \in \Bbbk[X]$ un polinomio mónico de grado n. El determinante de la matriz $X \cdot I_n - C(p) \in \mathsf{M}_n(\Bbbk(X))$ es*

$$\det(X \cdot I_n - C(p)) = p(X).$$

Demostración. Demostraremos la proposición haciendo inducción en n, observando que cuando $n = 1$ el resultado es inmediato.

Supongamos entonces que $n \geq 2$. Tenemos que calcular el determinante de la matriz

$$X \cdot I_n - C(p) = \begin{pmatrix} X & 0 & \dots & 0 & c_0 \\ -1 & X & \dots & 0 & c_1 \\ 0 & -1 & \dots & 0 & c_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \dots & -1 & X - c_{n-1} \end{pmatrix}.$$

Desarrollándolo a lo largo de la primera fila, vemos que

$$\begin{aligned} \det(X \cdot I_n - C(p)) &= \begin{vmatrix} X & 0 & \dots & 0 & c_0 \\ -1 & X & \dots & 0 & c_1 \\ 0 & -1 & \dots & 0 & c_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \dots & -1 & X - c_{n-1} \end{vmatrix} \\ &= X \begin{vmatrix} X & 0 & \dots & 0 & c_1 \\ -1 & X & \dots & 0 & c_2 \\ 0 & -1 & \dots & 0 & c_3 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \dots & -1 & X - c_{n-1} \end{vmatrix} + (-1)^{n+1} c_0 \begin{vmatrix} -1 & X & \dots & 0 \\ 0 & -1 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & -1 \end{vmatrix}. \end{aligned} \tag{16}$$

El primer determinante que aparece en el último miembto de esta igualdad es el de la matriz $X \cdot I_{n-1} - C(q) \in \mathsf{M}_n(\Bbbk(X))$, con q el polinomio

$$c_1 + c_2 X + \cdots + c_{n-1}X^{n-2} + X^{n-1} \in \Bbbk(X),$$

que es mónico y de grado $n-1$, así que la hipótesis inductiva nos dice que es igual a $q(X)$. El segundo determinante , por otro lado, es el de una matriz triangular inferior y de la Proposición 4.9.1 sabemos que vale $(-1)^{n-1}$. Volviendo con todo esto a la igualdad (16) vemos que

$$\det(X \cdot I_n - C(p)) = Xq(X) + c_0 = p(X),$$

como queremos. □

Matrices de Vandermonde

4.9.4. Si $n \in \mathbb{N}$ y $\alpha_1, \dots, \alpha_n \in \Bbbk$ son escalares, la ***matriz de Vandermonde*** para $\alpha_1, \dots, \alpha_n$ es la matriz $V(\alpha_1, \dots, \alpha_n) = (a_{i,j} \in \mathsf{M}_n(\Bbbk))$ que tiene $a_{i,j} = \alpha_j^{i-1}$ para cada $i, j \in [\![n]\!]$.

$$V(\alpha_1, \dots, \alpha_n) = \begin{pmatrix} 1 & 1 & \cdots & 1 \\ \alpha_1 & \alpha_2 & \cdots & \alpha_n \\ \vdots & \vdots & \ddots & \vdots \\ \alpha_1^{n-1} & \alpha_2^{n-1} & \cdots & \alpha_n^{n-1} \end{pmatrix}$$

El nombre recuerda a Alexandre-Théophile Vandermonde (1735–1796, Francia), a pesar de que no hay ningún registro de que lo haya considerado; ver, por ejemplo, el artículo [Yca13] para una discusión de esto. Vandermonde, sin embargo, es considerado el fundador de la teoría moderna de los determinantes: por ejemplo, es el primero en haber observado el efecto que tiene sobre el determinante de una matriz intercambiar dos de las columnas de ésta y que el determinante es es en consecuencia nulo cuando dos de ellas son iguales.

Proposición. *Sea $n \in \mathbb{N}$. Si $\alpha_1, \dots, \alpha_n \in \Bbbk$, entonces*

$$\det V(\alpha_1, \dots, \alpha_n) = \prod_{1 \le i < j \le n} (\alpha_i - \alpha_j) \tag{17}$$

Demostración. Hagamos inducción en n, observando que si $n = 1$ no hay nada que probar.

Escribamos V en lugar de $V(\alpha_1, \dots, \alpha_n)$, por simplicidad. Es claro que si existen i y j en $[\![n]\!]$ distintos tales que $\alpha_i = \alpha_j$, entonces $\det V(\alpha_1, \dots, \alpha_n) = 0$, ya que en ese caso la matriz tiene dos columnas iguales, y vale evidentemente la igualdad (17). Queda entonces considerar el caso en el que los escalares $\alpha_1, \dots, \alpha_n$ son distintos dos a dos.

Si todos los escalares $\alpha_1, \dots, \alpha_n$ son no nulos pongamos $k = 1$ y si no sea $k \in [\![n]\!]$ tal que $\alpha_k = 0$. Según la Proposición 4.6.2*(ii)*, desarrollando el determinante de V a lo largo de su columna k-ésima, es

$$\det V = \sum_{l=1}^{n} (-1)^{l+k} \alpha_k^{l-1} \det V^{(l,k)}.$$

Consideremos el polinomio

$$p(X) = \sum_{l=1}^{n} (-1)^{l+k} X^{l-1} \det V^{(l,k)} \in \Bbbk[X].$$

El grado de p es a lo sumo $n-1$ y el coeficiente de X^{n-1} en p es $(-1)^{n+k} \det V^{(n,k)}$. Como $V^{(n,k)}$ es la matriz de Vandermonde $V(\alpha_1, \dots, \widehat{\alpha_k}, \dots, \alpha_n)$, usando la hipótesis inductiva tenemos que

$$\det V^{(n,k)} = \det V(\alpha_1, \dots, \widehat{\alpha_k}, \dots, \alpha_n) = \prod_{\substack{1 \le i < j \le n \\ i,j \ne k}} (\alpha_i - \alpha_j) \ne 0.$$

Concluimos de esta forma que el polinomio p tiene grado exactamente $n-1$.

Si $i \in \{1, \dots, \widehat{k}, \dots, n\}$, entonces

$$p(\alpha_i) = \sum_{l=1}^{n} (-1)^{l+k} \alpha_i^{l-1} \det V^{(l,1)}$$

y esto es precisamente el valor del determinante de la matriz

$$V(\alpha_1, \dots, \underbrace{\alpha_i}_{k}, \dots, \alpha_n),$$

que tiene dos columnas iguales, la i-ésima y la k-ésima: vemos así que $p(\alpha_i) = 0$. El polinomio p, que tiene grado $n-1$, tiene entonces a los $n-1$ escalares $\alpha_1, \dots, \widehat{\alpha_k}, \dots, \alpha_n$, que son distintos dos a dos, como raíces y, en consecuencia, existe un escalar $\beta \in \Bbbk$ tal que

$$p(X) = \beta(X - \alpha_1) \cdots \widehat{(X - \alpha_k)} \cdots (X - \alpha_n).$$

De esta igualdad se sigue, en particular, que

$$p(0) = (-1)^{n-1} \beta \alpha_1 \cdots \widehat{\alpha_k} \cdots \alpha_n. \tag{18}$$

Por otro lado, de la definición de p es inmediato que

$$p(0) = (-1)^{k+1} \det V^{(1,k)} = (-1)^{k+1} \begin{vmatrix} \alpha_1 & \cdots & \alpha_{k-1} & \alpha_{k+1} & \cdots & \alpha_n \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ \alpha_1^{n-1} & \cdots & \alpha_{k-1}^{n-1} & \alpha_{k+1}^{n-1} & \cdots & \alpha_n^{n-1} \end{vmatrix} \tag{19}$$

y, usando la homogeneidad del determinante de una matriz con respecto a las columnas de ésta, este último determinante es

$$\begin{aligned} \begin{vmatrix} \alpha_1 & \cdots & \alpha_{k-1} & \alpha_{k+1} & \cdots & \alpha_n \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ \alpha_1^{n-1} & \cdots & \alpha_{k-1}^{n-1} & \alpha_{k+1}^{n-1} & \cdots & \alpha_n^{n-1} \end{vmatrix} &= \alpha_1 \cdots \widehat{\alpha_k} \cdots \alpha_n \begin{vmatrix} 1 & \cdots & 1 & 1 & \cdots & 1 \\ \alpha_1 & \cdots & \alpha_{k-1} & \alpha_{k+1} & \cdots & \alpha_n \\ \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ \alpha_1^{n-2} & \cdots & \alpha_{k-1}^{n-2} & \alpha_{k+1}^{n-2} & \cdots & \alpha_n^{n-2} \end{vmatrix} \\ &= \alpha_1 \cdots \widehat{\alpha_k} \cdots \alpha_n \det V(\alpha_1, \dots, \widehat{\alpha_k}, \dots, \alpha_n). \end{aligned}$$

Comparando las expresiones (18) y (19) para $p(0)$ concluimos entonces que

$$(-1)^{n-1} \beta \alpha_1 \cdots \widehat{\alpha_k} \cdots \alpha_n = (-1)^k \alpha_1 \cdots \widehat{\alpha_k} \cdots \alpha_n \det V(\alpha_1, \dots, \widehat{\alpha_k}, \dots, \alpha_n)$$

Como $\alpha_1 \cdots \widehat{\alpha_k} \cdots \alpha_n \neq 0$, esto nos dice que

$$\beta = (-1)^{n-k-1} \det V(\alpha_1, \dots, \widehat{\alpha_k}, \dots, \alpha_n) = (-1)^{n-k-1} \prod_{\substack{1 \leq i < j \leq n \\ i,j \neq k}} (\alpha_i - \alpha_j),$$

usando en la última igualdad otra vez la hipótesis inductiva y, en definitiva, que

$$p(X) = (-1)^{n-k-1} \prod_{\substack{1 \leq i < j \leq n \\ i,j \neq k}} (\alpha_i - \alpha_j) \cdot (X - \alpha_1) \cdots \widehat{(X - \alpha_k)} \cdots (X - \alpha_n).$$

En particular, evaluando esto en α_k vemos que

$$p(\alpha_k) = (-1)^{n-k-1} \prod_{\substack{1\le i<j\le n \\ i,j\neq k}} (\alpha_i - \alpha_j) \cdot (\alpha_k - \alpha_1)\cdots\widehat{(\alpha_k - \alpha_k)}\cdots(\alpha_k - \alpha_n) = \prod_{1\le i<j\le n} (\alpha_i - \alpha_j).$$

Esto completa la inducción, ya que $p(\alpha_k) = \det V(\alpha_1, \dots, \alpha_k)$. □

§10. El determinante de un endomorfismo

4.10.1. Si V es un espacio vectorial de dimensión finita, $\mathscr{B}$ una base ordenada de V y $f : V \to V$ un endomorfismo de V, llamamos ***determinante de f*** al determinante de la matriz $[f]^{\mathscr{B}}_{\mathscr{B}}$. Esta definición tiene sentido porque este escalar no depende de la elección de la base $\mathscr{B}$ sino solamente de f: esto es el contenido de la siguiente proposición.

Proposición. *Sea V un espacio vectorial de dimensión finita y sean $\mathscr{B}$ y $\mathscr{B}'$ dos bases ordenadas de V. Si $f : V \to V$ es un endomorfismo de V, entonces*

$$\det[f]^{\mathscr{B}}_{\mathscr{B}} = \det[f]^{\mathscr{B}'}_{\mathscr{B}'}.$$

Demostración. Sabemos de la Proposición 2.8.7 que

$$[f]^{\mathscr{B}'}_{\mathscr{B}'} = C(\mathscr{B}, \mathscr{B}') \cdot [f]^{\mathscr{B}}_{\mathscr{B}} \cdot C(\mathscr{B}', \mathscr{B})$$

y de acuerdo a la Proposición 1.8.4*(iii)* las matrices $C(\mathscr{B}, \mathscr{B}')$ y $C(\mathscr{B}, \mathscr{B}')$ son mutuamente inversas, así que el Corolario 4.3.4 nos permite concluir lo que queremos. □

4.10.2. Proposición. *Sea V un espacio vectorial de dimensión finita.*

(*i*) *Si $\mathrm{id}_V : V \to V$ es el endomorfismo identidad de V, entonces $\det \mathrm{id}_V = 1$.*

(*ii*) *Si $f, g : V \to V$ son dos endomorfismos de V, entonces $\det(f \circ g) = \det f \cdot \det g$.*

(*iii*) *Si $h : V \to V$ es un endomorfismo de V, entonces h es un automorfismo exactamente cuando $\det h \neq 0$ y en ese caso se tiene que $\det h^{-1} = (\det h)^{-1}$.*

Demostración. Sea $n = \dim V$ y sea $\mathscr{B}$ una base ordenada de V. Sabemos que $[\mathrm{id}_V] = I_n$, la matriz identidad de $\mathsf{M}_n(\Bbbk)$, así que que $\det \mathrm{id}_V = \det I_n = 1$: esto prueba (*i*).

Si $f, g : V \to V$ son endomorfismos de V, entonces

$$\det(f \circ g) = \det[f \circ g]^{\mathscr{B}}_{\mathscr{B}} = \det\big([f]^{\mathscr{B}}_{\mathscr{B}} \cdot [g]^{\mathscr{B}}_{\mathscr{B}}\big) = \det[f]^{\mathscr{B}}_{\mathscr{B}} \cdot \det[g]^{\mathscr{B}}_{\mathscr{B}} = \det f \cdot \det g,$$

como afirma (*ii*). Por otro lado, si $h : V \to V$ es un endomorfismo, sabemos de la Proposición 2.8.6 que h es un isomofismo si y solamente si la matriz $[h]^{\mathscr{B}}_{\mathscr{B}}$ es inversible y, de acuerdo a la Proposición 4.7.1*(ii)*, esto ocurre exactamente cuando el escalar $\det h = \det[h]^{\mathscr{B}}_{\mathscr{B}}$ es no nulo. Si ése es el caso, entonces

$$\det h \cdot \det h^{-1} = \det(h \circ h^{-1}) = \det \mathrm{id}_V = 1$$

y, por supuesto, se sigue que $\det h^{-1} = (\det h)^{-1}$. □

4.10.3. Proposición. *Sea V un espacio vectorial de dimensión finita y sea $f : V \to V$ un endomorfismo de V. Si W es un subespacio f-invariante de V, entonces*

$$\det f = \det f_W \cdot \det f^W.$$

Demostración. Sean n y r las dimensiones de V y de W y sea $\mathscr{B} = (x_1, \dots, x_n)$ una base de V tal que $\mathscr{B}' = (x_1, \dots, x_r)$ es una base ordenada de W. De acuerdo a la Proposición 2.12.5, entonces, $\mathscr{B}'' = ([x_{r+1}], \dots, [x_n])$ es una base ordenada de V/W y hay una matriz $C \in \mathsf{M}_{r,n-r}(\Bbbk)$ tal que la matriz de f con respecto a $\mathscr{B}$ es triangular superior por bloques de la forma

$$[f]^{\mathscr{B}}_{\mathscr{B}} = \begin{pmatrix} [f_W]^{\mathscr{B}'}_{\mathscr{B}'} & C \\ 0 & [f^W]^{\mathscr{B}''}_{\mathscr{B}''} \end{pmatrix}$$

La conclusión de la proposición es consecuencia de esto y de la Proposición 4.9.2. □

Capítulo 5

Autovectores y autovalores

§1. Autovectores y autovalores

5.1.1. Sea V un espacio vectorial y sea $f : V \to V$ un endomorfismo de V. Un vector no nulo x de V es un ***autovector*** de f si existe un escalar $\lambda \in \Bbbk$ tal que $f(x) = \lambda x$ y en ese caso decimos que λ es el ***autovalor*** correspondiente a x o que x es un autovector de autovalor λ. Por otro lado, un escalar $\lambda \in \Bbbk$ es un ***autovalor*** de f si existe un autovector de f que tenga a λ por autovalor. El ***espectro*** de f es el conjunto $\mathsf{Spec}(f)$ de sus autovalores.

5.1.2. Ejemplos.

(a) Si V es un espacio vectorial y $\lambda \in \Bbbk$ es un escalar, entonces todo vector no nulo de V es un autovector de la función lineal $f : x \in V \mapsto \lambda x \in V$ y el escalar λ es el único autovalor de f.

(b) Si V es un espacio vectorial y $f : V \to V$ es un endomorfismo, los autovectores de f que tienen a 0 como autovalor son precisamente los elementos no nulos del núcleo $\mathsf{Nu}(f)$.

(c) Consideremos la función lineal $f : (x, y) \in \mathbb{R}^2 \mapsto (y, -x) \in \mathbb{R}^2$, endomorfismo del espacio vectorial real $\mathbb{R}^2$. Si $v = (x, y)$ es un elemento no nulo de $\mathbb{R}^2$, entonces es fácil ver que v y $f(v)$ son linealmente independientes: esto implica, en particular, que $f(v)$ no es un múltiplo escalar de v y, entonces, que v no es un autovector de f. Vemos así que esta función lineal no posee ningún autovector ni ningún autovalor.

En cambio, el endomorfismo $(x, y) \in \mathbb{C}^2 \mapsto (y, -x) \in \mathbb{C}^2$ del espacio vectorial complejo $\mathbb{C}^2$ tiene a los vectores $(1, i)$ y $(1, -i)$ como autovectores, con autovalores correspondientes los escalares i y $-i$.

(d) Sea $C^\infty(\mathbb{R})$ el espacio vectorial real de las funciones $\mathbb{R} \to \mathbb{R}$ que son derivables infinitas veces y sea $D : f \in C^\infty(\mathbb{R}) \to f' \in C^\infty(\mathbb{R})$. Supongamos que $f \in C^\infty(\mathbb{R})$ es un autovector de D de autovalor $\lambda \in \mathbb{R}$, de manera que

$$f' = \lambda f. \tag{1}$$

La función $g : t \in \mathbb{R} \mapsto e^{-\lambda t} f(t) \in \mathbb{R}$ es claramente derivable con continuidad en todo su dominio y tiene en cada $t \in \mathbb{R}$ derivada

$$g'(t) = -\lambda e^{-\lambda t} f(t) + e^{-\lambda t} f'(t) = 0,$$

en vista de la relación (1). Esto implica que la función g es constante y, como $g(0) = f(0)$, que $e^{-\lambda t} f(t) = f(0)$ para todo $t \in \mathbb{R}$, esto es, que $f(t) = f(0)e^{-\lambda t}$. Vemos así que f es un múltiplo escalar de la función exponencial

$$h_\lambda : x \in \mathbb{R} \mapsto e^{-\lambda t} \in \mathbb{R}.$$

Un cálculo directo muestra que esta función h_λ es en efecto un autovector de D de autovalor λ, junto con todos sus múltiplos escalares, así que podemos concluir con esto que

> *para cada $\lambda \in \mathbb{R}$ los autovectores de $D : f \in C^\infty(\mathbb{R}) \to f' \in C^\infty(\mathbb{R})$ de autovalor λ son precisamente los múltiplos escalares no nulos de la función h_λ.*

En particular, todo elemento de $\mathbb{R}$ es un autovalor de esta función D.

(e) Podemos considerar también la función dada por la derivación $D : p \in \mathbb{R}[X] \mapsto p' \in \mathbb{R}[X]$ pero ahora definida sobre el espacio vectorial $\mathbb{R}[X]$ de los polinomios con coeficientes reales.

Supongamos que $p \in \mathbb{R}[X]$ es un autovector de D de autovalor λ, de manera que $p' = \lambda p$. Como p no es nulo, tiene un grado d bien definido. Si $d > 0$ y $\lambda \neq 0$, entonces λp tiene grado d y p' tiene grado $d - 1$: esto es imposible, ya que estos dos polinomios son iguales. Por otro lado, si $\lambda = 0$, entonces p' es el polinomio nulo y que, por lo tanto, p es constante. Vemos así que tiene que ser $d = 0$, que el polinomio p es constante y que $\lambda = 0$.

Hemos mostrado, en definitiva, que

> *el único autovalor de $D : p \in \mathbb{R}[X] \mapsto p' \in \mathbb{R}[X]$ es $\lambda = 0$, y los autovectores correspondientes son los polinomios constantes no nulos.*

Vemos así que, a diferencia de lo que sucede en el ejemplo anterior, el único autovalor de la función lineal D es 0. ◇

5.1.3. Una observación sencilla es que el conjunto de autovalores de un endomorfismo de un espacio vectorial correspondientes a un autovalor fijo es un subespacio, una vez que uno agrega el vector nulo. Esto es consecuencia de la siguiente proposición.

Proposición. *Sea V un espacio vectorial y sea $f : V \to V$ un endomorfismo de V. Para cada $\lambda \in \Bbbk$, el subconjunto $E_\lambda(f) = \{x \in V : f(x) = \lambda x\}$ de V es un subespacio de V.*

Llamamos al subespacio $E_\lambda(f)$ el ***autoespacio*** de f correspondiente a λ. Si tiene dimensión finita, llamamos al número $\dim E_\lambda(f)$ la ***multiplicidad geométrica*** de λ como autovalor de f. Es claro que esta multiplicidad es positiva si y solamente si λ es un autovalor de f.

Demostración. La afirmación de la proposición es consecuencia de que $E_\lambda(f)$ es el núcleo de la función lineal $\lambda \mathrm{id}_V - f : V \to V$. □

5.1.4. Una segunda observación importante es que autovectores correspondientes a autovalores distintos dos a dos son linealmente independientes.

Proposición. *Sea V un espacio vectorial y sea $f : V \to V$ un endomorfismo de V. Si $x_1, \dots, x_n \in V$ son autovectores de f correspondientes a autovalores $\lambda_1, \dots, \lambda_n \in \Bbbk$ y estos últimos son distintos dos a dos, entonces los vectores $x_1, \dots, x_n$ son linealmente independientes.*

Demostración. Sean $x_1, \dots, x_n \in V$ autovectores de f correspondientes a autovalores $\lambda_1, \dots, \lambda_n \in \Bbbk$ distintos dos a dos y sean $a_1, \dots, a_n \in \Bbbk$ escalares tales que $a_1 x_1 + \cdots + a_n x_n = 0$. Para cada $j \in [\![n]\!]$ podemos aplicar la función f^{j-1} a ambos lados de esta igualdad y ver que

$$\sum_{k=1}^{n} a_k \lambda_k^{j-1} x_k = 0. \tag{2}$$

Ahora bien, como los escalares $\lambda_1, \dots, \lambda_n$ son distintos dos a dos, es una consecuencia inmediata de la Proposición 4.9.4 que la matriz de Vandermonde

$$V(\lambda_1, \dots, \lambda_n) = \begin{pmatrix} 1 & 1 & \cdots & 1 \\ \lambda_1 & \lambda_2 & \cdots & \lambda_n \\ \vdots & \vdots & \ddots & \vdots \\ \lambda_1^{n-1} & \lambda_2^{n-1} & \cdots & \lambda_n^{n-1} \end{pmatrix}$$

es inversible. Supongamos que su matriz inversa es $V(\lambda_1, \dots, \lambda_n)^{-1} = (\mu_{i,j}) \in \mathsf{M}_n(\Bbbk)$, de manera que, en particular, para cada $i, k \in [\![n]\!]$ se tiene que

$$\sum_{j=1}^{n} \mu_{i,j} \lambda_k^{j-1} = \begin{cases} 1, & \text{si } i = j; \\ 0, & \text{en caso contrario.} \end{cases} \tag{3}$$

Para cada $i \in [\![n]\!]$, vemos entonces —usando primero la ecuación (2) y luego la (3)— que

$$0 = \sum_{j=1}^{n} \mu_{i,j} \left(\sum_{k=1}^{n} a_k \lambda_k^{j-1} x_k \right) = \sum_{k=1}^{n} a_k \left(\sum_{j=1}^{n} \mu_{i,j} \lambda_k^{j-1} \right) x_k = a_i x_i$$

y, en consecuencia, que los escalares $a_1, \dots, a_n$ son todos nulos. Esto prueba la proposición. □

5.1.5. El siguiente corolario de la Proposición 5.1.4 es inmediato:

Corolario. *Sea V un espacio vectorial y sea $f : V \to V$ un endomorfismo de V. Si V tiene dimensión finita, entonces f posee un numero finito de autovalores y, de hecho, el número de éstos es a lo sumo igual a* $\dim V$.

Demostración. En efecto, si $\lambda_1, \dots, \lambda_n$ son autovalores de f distintos dos a dos, existen vectores no nulos $x_1, \dots, x_n \in V$ tales que $f(x_i) = \lambda_i x_i$ para cada $i \in [\![n]\!]$ y la Proposición 5.1.4 implica que estos n vectores son linealmente independientes. Por supuesto, de esto se deduce inmediatamente que $n \leq \dim V$. □

5.1.6. La Proposición 5.1.4 tiene, por otro lado, la siguiente consecuencia importante:

Corolario. *Sea V un espacio vectorial y sea : $V \to V$ un endomorfismo de V. Si $\lambda_1, \dots, \lambda_n \in \Bbbk$ son escalares distintos dos a dos, entonces los subespacios $E_{\lambda_1}(f), \dots, E_{\lambda_n}(f)$ de V son independientes.*

Demostración. Podemos usar la caracterización de la independencia de subespacios que da la tercera condición de la Proposición 1.10.3: basta mostrar que si $x_1, \dots, x_n$ son vectores de V tales que para cada $i \in [\![n]\!]$ es $x_i \in E_{\lambda_i}(f)$ y se tiene que $x_1 + \cdots + x_n = 0$, entonces cada uno de los vectores $x_1, \dots, x_n$ es nulo. Eso sigue inmediatamente de la Proposición 5.1.4. □

5.1.7. Sea $n \in \mathbb{N}$, sea $A \in \mathsf{M}_n(\Bbbk)$ una matriz cuadrada de tamaño n con entradas en $\Bbbk$ y consideremos el endomorfismo $f_A : x \in \Bbbk^n \mapsto Ax \in \Bbbk^n$ de $\Bbbk^n$. Aprovechando la estrecha relación que hay entre la matriz A y la función f_A, podemos adaptar las nociones presentadas en esta sección a la matriz A de la siguiente manera:

- Si $x \in \Bbbk^n$ y $\lambda \in \Bbbk$, decimos que x es un ***autovector de A de autovalor λ*** si x es un autovector de f_A de autovalor λ.
- Un escalar $\lambda \in \Bbbk$ es un ***autovalor de A*** si la matriz A posee un autovector de autovalor λ, y el ***espectro*** de A es el conjunto $\mathsf{Spec}(A)$ de sus autovalores.
- Si $\lambda \in \Bbbk$, entonces el ***autoespacio de A correspondiente a λ*** es el subespacio $E_\lambda(f_A)$, que escribimos más simplemente en la forma $E_\lambda(A)$.

Es inmediato, en vista de la forma de estas definiciones, que valen para matrices y sus autovectores y autovalores propiedades análogas a las enunciadas arriba para endomorfismos: en todos los casos, estos resultados se obtienen para la matriz A de considerar el resultado correspondiente para el endomorfismo f_A. Esto mismo puede decirse de casi todos los enunciados que siguen.

5.1.8. Ejemplo. Sea $n \in \mathbb{N}$, sea $p(X) = c_0 + c_1X + \cdots + c_{n-1}X^{-1} + X^n \in \Bbbk[X]$ un polinomio mónico de grado n y sea $C(p)$ la matriz compañera de p,

$$C(p) = \begin{pmatrix} 0 & 0 & \dots & 0 & -c_0 \\ 1 & 0 & \dots & 0 & -c_1 \\ 0 & 1 & \dots & 0 & -c_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \dots & 1 & -c_{n-1} \end{pmatrix} \in \mathsf{M}_n(\Bbbk),$$

Si $\lambda \in \Bbbk$ es una raíz de p, de manera que $p(\lambda) = 0$, un cálculo directo muestra que el vector $x_\lambda = (1, \lambda, \dots, \lambda^{n-1})^{\mathrm{t}} \in \Bbbk^n$ es tal que $C(p)x_\lambda = \lambda x_\lambda$, es decir, que se trata de un autovector de $C(p)$ de autovalor λ. Vemos así que el espectro de la matriz $C(p)$ contiene al conjunto de raíces de p en $\Bbbk$. ◇

§2. Diagonalizabilidad

5.2.1. Decimos que un endomorfismo $f : V \to V$ de un espacio vectorial V es ***diagonalizable*** si existe una base $\mathscr{B}$ de V cuyos elementos son autovectores de f. La motivación para la elección de este adjetivo viene del siguiente resultado:

Proposición. *Sea V un espacio vectorial de dimensión finita y sea $f : V \to V$ un endomorfismo de V. Las siguientes afirmaciones son equivalentes:*

(a) *El endomorfismo f es diagonalizable.*

(b) *Existe una base ordenada $\mathscr{B}$ de V tal que la matriz $[f]^{\mathscr{B}}_{\mathscr{B}}$ es diagonal.*

(c) *Existen escalares $\mu_1, \dots, \mu_r \in \Bbbk$ distintos dos a dos tales que $V = E_{\mu_1}(f) \oplus \cdots \oplus E_{\mu_r}(f)$.*

Hacer: Agregar: f es combinación lineal de un conjunto de idempotentes ortogonales.

Demostración. Sea n la dimensión de V.

$(a) \Rightarrow (b)$ Supongamos que el endomorfismo f es diagonalizable y sea $\mathscr{B} = (x_1, \dots, x_n)$ una base ordenada de V cuyos elementos son autovectores de f, de manera que existen escalares $\lambda_1, \dots, \lambda_n \in \Bbbk$ tales que para cada $i \in [\![n]\!]$ se tiene que $f(x_i) = \lambda_i x_i$. Si $[f]^{\mathscr{B}}_{\mathscr{B}} = (a_{i,j}) \in \mathsf{M}_n(\Bbbk)$ es la matriz de f con respecto a la base $\mathscr{B}$, vale entonces que

$$a_{i,j} = \begin{cases} \lambda_i, & \text{si } i = j; \\ 0, & \text{en caso contrario.} \end{cases}$$

Esto, por supuesto, nos dice que $[f]^{\mathscr{B}}_{\mathscr{B}}$ es una matriz diagonal.

$(b) \Rightarrow (c)$ Supongamos ahora que $\mathscr{B} = (x_1, \dots, x_n)$ es una base ordenada de V tal que la matriz $[f]^{\mathscr{B}}_{\mathscr{B}}$ es diagonal y sean $\lambda_1, \dots, \lambda_n \in \Bbbk$ los escalares que aparecen en su diagonal, en orden. Estos n escalares no son necesariamente distintos: sean $\mu_1, \dots, \mu_r$ los elementos listados sin repeticiones del conjunto $\{\lambda_1, \dots, \lambda_n\}$, de manera que hay una función $\phi : [\![n]\!] \to [\![r]\!]$ tal que $\lambda_i = \mu_{\phi(i)}$ para cada $i \in [\![n]\!]$. Del Corolario 5.1.6 sabemos que los subespacios $E_{\mu_1}(f), \dots, E_{\mu_r}(f)$ son independientes, así que si llamamos V' a su suma, tenemos que $V' = E_{\mu_1}(f) \oplus \cdots \oplus E_{\mu_r}(f)$. Para cada $i \in [\![n]\!]$ es $f(x_i) = \lambda_i x_i = \mu_{\phi(i)} x_i$, así que $x_i \in E_{\mu_{\phi(i)}}(f)$. Vemos de esta forma que $\mathscr{B} \subseteq V'$ y, por lo tanto, que $V = \langle \mathscr{B} \rangle \subseteq V'$: esto nos dice, claro, que $V = V'$ y prueba que vale (c).

$(c) \Rightarrow (a)$ Supongamos que existen escalares $\mu_1, \dots, \mu_r \in \Bbbk$ distintos dos a dos tales que $V = E_{\mu_1}(f) \oplus \cdots \oplus E_{\mu_r}(f)$. Para cada $i \in [\![r]\!]$ sea $\mathscr{B}_i$ una base de $E_{\mu_i}(f)$. De la Proposición 1.10.4 sabemos que la unión $\mathscr{B} = \mathscr{B}_1 \cup \cdots \cup \mathscr{B}_r$ es una base de V. Para terminar, basta ver que todos sus elementos son autovectores de f: esto es claro, ya que para cada $i \in [\![r]\!]$ los elementos de $\mathscr{B}_i$ son autovectores de f de autovalor μ_i porque están en $E_{\mu_i}(f)$ y no son nulos. □

5.2.2. Si $n \in \mathbb{N}$ y $A \in \mathsf{M}_n(\Bbbk)$ es una matriz cuadrada de tamaño n, decimos que A es ***diagonalizable*** si hay una base de $\Bbbk^n$ cuyos elementos son autovectores de A.

Proposición. *Sea $n \in \mathbb{N}$ y sea $A \in \mathsf{M}_n(\Bbbk)$. Las siguientes afirmaciones son equivalentes:*

(a) *La matriz A es diagonalizable.*

(b) *El endomorfismo $f_A : x \in \Bbbk^n \mapsto Ax \in \Bbbk^n$ de $\Bbbk^n$ es diagonalizable.*

(c) *Existe una matriz inversible $C \in \mathsf{GL}_n(\Bbbk)$ tal que la matriz $C^{-1}AC$ es diagonal.*

Si estas afirmaciones se cumplen y C es una matriz como en (c)*, entonces el conjunto de los vectores de $\Bbbk^n$ dados por las columnas de C es una base de $\Bbbk^n$ formada por autovectores de A.*

Demostración. Como los autovectores de la matriz A y los autovectores del endomorfismo f_A coinciden, la equivalencia $(a) \Leftrightarrow (b)$ es inmediata.

Sea $\mathscr{B}$ la base ordenada estándar de $\Bbbk^n$, sea $\mathscr{B}'$ una base ordenada cualquiera de $\Bbbk^n$ y sea $C = C(\mathscr{B}', \mathscr{B})$ la matriz de cambio de base de $\mathscr{B}'$ a $\mathscr{B}$. Sabemos que C es inversible, que $[f_A]_{\mathscr{B}}^{\mathscr{B}} = A$ y, según la Proposición 2.8.7, que

$$[f_A]_{\mathscr{B}'}^{\mathscr{B}'} = C(\mathscr{B}, \mathscr{B}') \cdot [f_A]_{\mathscr{B}}^{\mathscr{B}} \cdot C(\mathscr{B}', \mathscr{B}) = C^{-1}AC. \tag{4}$$

Si f_A es diagonalizable, entonces eligiendo en lo anterior a la base ordenada $\mathscr{B}'$ de manera tal que la matriz $[f_A]_{\mathscr{B}}^{\mathscr{B}'}$ sea diagonal, vemos que $C^{-1}AC$ es diagonal. Esto prueba que $(b) \Rightarrow (c)$.

Recíprocamente, supongamos que vale (c), sea $C \in \mathsf{GL}_n(\Bbbk)$ una matriz inversible tal que $C^{-1}AC$ es diagonal y elijamos en lo anterior como $\mathscr{B}'$ a la base ordenada de $\Bbbk^n$ tal que $C(\mathscr{B}', \mathscr{B}) = C$, cuya existencia garantiza la Proposición 1.8.5: la igualdad (4) nos dice entonces que la matriz $[f]_{\mathscr{B}'}^{\mathscr{B}'}$ es diagonal y, por lo tanto, que el endomorfismo f_A es diagonalizable, como afirma (b). Vemos de esta manera que $(c) \Rightarrow (b)$, lo que completa la demostración de la equivalencia de las tres afirmaciones del enunciado.

Supongamos ahora que estas afirmaciones se cumplen y sea $C \in \mathsf{GL}_n(\Bbbk)$ una matriz inversible tal que $D = C^{-1}AC$ es diagonal. Para cada $i \in [\![n]\!]$ sea x_i el elemento de $\Bbbk^n$ dado por la i-ésima columna de C y sea λ_i la entrada (i,i)-ésima de D. Como la matriz C es inversible, la Proposición 4.8.1 nos dice que $(x_1, \dots, x_n)$ es una base ordenada de V. Por otro lado, si $(e_1, \dots, e_n)$ es la base ordenada estándar de $\Bbbk^n$, para cada $i \in [\![n]\!]$ vale que $x_i = Ce_i$ y entonces

$$Ax_i = ACe_i = C^{-1}De_i = \lambda_i C^{-1}e_i = \lambda_i x_i.$$

Los elementos de la base $(x_1, \dots, x_n)$ son por lo tanto autovectores de A. □

5.2.3. Ejemplo. Sea $n \in \mathbb{N}$, sea $p \in \Bbbk[X]$ un polinomio mónico de grado n y sea $C(p)$ la matriz compañera de p. En el Ejemplo 5.2.3 vimos que cada vez que $\lambda \in \Bbbk$ es una raíz de p el vector $x_\lambda = (1, \lambda, \dots, \lambda^{n-1})^{\mathsf{t}} \in \Bbbk^n$ es un autovector de $C(p)$ de autovalor λ. Se sigue de esto que si p posee n raíces $\lambda_1, \dots, \lambda_n$ en $\Bbbk$ distintas dos a dos, entonces el conjunto $\{x_{\lambda_1}, \dots, x_{\lambda_n}\}$ es linealmente independiente —de acuerdo a la Proposición 5.1.4— y, como tiene cardinal igual a la dimensión de $\Bbbk^n$, es una base de este espacio vectorial: la matriz $C(p)$ es por lo tanto diagonalizable. La matriz $C \in M_n(\Bbbk)$ cuyas columnas son los vectores $x_{\lambda_1}, \dots, x_{\lambda_n}$, en ese orden, es precisamente la matriz de Vandermonde $V(\lambda_1, \dots, \lambda_n)$ de 4.9.4, y se sigue de la Proposición 5.2.2 que el producto $C^{-1} \cdot C(p) \cdot C$ es una matriz diagonal que tiene a lo largo de la diagonal a los escalares $\lambda_1, \dots, \lambda_n$, en ese orden. ◇

5.2.4. Es importante tener en mente que no todo endomorfismo ni toda matriz es diagonalizable.

Por ejemplo, supongamos que la matriz

$$A = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \in \mathsf{M}_2(\mathbb{k})$$

es diagonalizable, de manera que existen una matriz inversible $C \in \mathsf{M}_2(\mathbb{k})$ y escalares $\lambda_1, \lambda_2 \in \mathbb{k}$ tales que

$$C^{-1}AC = \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix}. \tag{5}$$

Como $A^2 = 0$, se tiene entonces que

$$0 = C^{-1}A^2C = C^{-1}AC \cdot C^{-1}AC = \begin{pmatrix} \lambda_1 & 0 \\ 0 & \lambda_2 \end{pmatrix}^2 = \begin{pmatrix} \lambda_1^2 & 0 \\ 0 & \lambda^2 \end{pmatrix}$$

y esto sólo es posible si $\lambda_1 = \lambda_2 = 0$. En vista de la igualdad (5), esto implica que $C^{-1}AC = 0$, lo que nos dice que $A = 0$: esto es absurdo.

§3. El polinomio característico

5.3.1. Podemos caracterizar los autovalores de un endomorfismo de un espacio vectorial de dimensión finita o de una matriz de la siguiente manera:

Proposición.

(i) *Sea V un espacio vectorial de dimensión finita, sea $\mathscr{B}$ una base ordenada de V y sea $f : V \to V$ un endomorfismo de V. Si $\lambda \in \mathbb{k}$, las condiciones siguientes son equivalentes:*
   - (a) *El escalar λ es un autovalor de f.*
   - (b) *Es $\det(\lambda\,\mathrm{id}_V - f) = 0$.*

(ii) *Sea $n \in \mathbb{N}$ y sea $A \in \mathsf{M}_n(\mathbb{k})$. Si $\lambda \in \mathbb{k}$, las condiciones siguientes son equivalentes:*
   - (a) *El escalar λ es un autovalor de A.*
   - (b) *Es $\det(\lambda I_n - A) = 0$.*

Demostración. En vista de nuestras definiciones, es suficiente que probemos la primera parte de la proposición, ya que la segunda sigue inmediatamente de ésta y de la observación de que si $f_A : x \in \mathbb{k}^n \mapsto Ax \in \mathbb{k}^n$, entonces $\det(\lambda\mathrm{id}_{\mathbb{k}^n} - f_A) = \det(\lambda I_n - A)$.

Sea $\lambda \in \mathbb{k}$. Si λ es un autovalor de f, el subespacio $E_\lambda(f) = \mathsf{Nu}(\lambda\mathrm{id}_V - f)$ es no nulo y, en particular, el endomorfismo $\lambda\mathrm{id}_V - f : V \to V$ no es un isomorfismo. La Proposición 4.10.2*(iii)* nos dice entonces que el determinante de este endomorfismo es nulo. Recíprocamente, si vale la condición (b) del enunciado, el endomorfismo $\lambda\mathrm{id}_V - f$ no es un isomorfismo y, en vista de la Proposición 2.5.3, tampoco un monomorfismo. Su núcleo es por lo tanto no nulo y como éste coincide con $E_\lambda(f)$, esto nos dice que λ es un autovalor de f. □

5.3.2. La Proposición 5.3.1 nos lleva a hacer la siguiente definición. Si $n \in \mathbb{N}$ y $A \in \mathsf{M}_n(\Bbbk)$, podemos ver a A como un elemento de $M_n(\Bbbk(X))$ y considerar la matriz $XI_n - A$. Llamamos ***polinomio característico de** A* al polinomio

$$\chi_A = \det(XI_n - A).$$

Observemos que, en vista del Corolario 4.4.15*(ii)*, esto es efectivamente un elemento de $\Bbbk[X]$ y no solamente de $\Bbbk(X)$, ya todas las entradas de la matriz $XI_n - A$ están en $\Bbbk[X]$.

5.3.3. Ejemplos.

(a) Sea $n \in \mathbb{N}$ y sea $A = (a_{i,j}) \in \mathsf{M}_n(\Bbbk)$ una matriz triangular superior, de manera que $a_{i,j} = 0$ si $i, j \in [\![n]\!]$ son tales que $i > j$. Como la matriz $XI_n - A \in \mathsf{M}_n(\Bbbk(X))$ es ella también triangular superior, la Proposición 4.9.1 nos permite calcular inmediatamente su determinante y vemos que

$$\chi_A = \det(XI_n - A) = (X - a_{1,1}) \cdots (X - a_{n,n}).$$

(b) Más generalmente, sean $n, r \in \mathbb{N}$, sean $n_1, \dots, n_r \in \mathbb{N}$ tales que $n = n_1 + \cdots + n_r$ y supongamos que para cada $i, j \in [\![r]\!]$ con $i \leq j$ tenemos una matriz $A_{i,j} \in \mathsf{M}_{n_i,n_j}(\Bbbk)$. Si A es la matriz triangular superior por bloques

$$\begin{pmatrix} A_{1,1} & A_{1,2} & \cdots & A_{1,r-1} & A_{1,r} \\ 0 & A_{2,2} & \cdots & A_{2,r-1} & A_{2,r} \\ \vdots & & \ddots & \vdots & \vdots \\ 0 & & & A_{r-1,r-1} & A_{r-1,r} \\ 0 & 0 & \cdots & 0 & A_{r,r} \end{pmatrix},$$

entonces el polinomio característico de A es el producto de los polinomios característicos de las matrices que aparecen a lo largo de la diagonal, esto es,

$$\chi_A = \chi_{A_{1,1}} \cdots \chi_{A_{r,r}}.$$

Esto es consecuencia de la Proposición 4.9.2 y de que

$$XI_n - A = \begin{pmatrix} XI_{n_1} - A_{1,1} & -A_{1,2} & \cdots & -A_{1,r-1} & -A_{1,r} \\ 0 & XI_{n_2} - A_{2,2} & \cdots & -A_{2,r-1} & -A_{2,r} \\ \vdots & & \ddots & \vdots & \vdots \\ 0 & & \cdots & XI_{n_{r-1}} - A_{r-1,r-1} & -A_{r-1,r} \\ 0 & 0 & \cdots & 0 & XI_{n_r} - A_{r,r} \end{pmatrix}$$

también es una matriz triangular superior por bloques.

(c) Sea $n \in \mathbb{N}$, sea $p \in \Bbbk[X]$ un polinomio mónico de grado n y sea $C(p) \in \mathsf{M}_n(\Bbbk)$ la matriz compañera de p. En 4.9.3 calculamos el determinante de la matriz $XI_n - C(p)$ y ese cálculo significa, precisamente, que el polinomio característico de $C(p)$ es

$$\chi_{C(p)} = \det(XI_n - C(p)) = p.$$

(d) Si $n \in \mathbb{N}$ y $A \in \mathsf{M}_n(\Bbbk)$, entonces el polinomio característico de A y el de su transpuesta A^{t} son iguales. En efecto, como $XI_n - A^{\mathrm{t}} = (XI_n - A)^{\mathrm{t}}$, de la Proposición 4.5.1 vemos que

$$\chi_{A^{\mathrm{t}}} = \det(XI_n - A^{\mathrm{t}}) = \det(XI_n - A)^{\mathrm{t}} = \det(XI_n - A) = \chi_A. \qquad \diamond$$

5.3.4. Una observación importante es que dos matrices que son conjugadas tienen el mismo polinomio característico:

Proposición. *Sea $n \in \mathbb{N}$ y sean A y B dos matrices de $\mathsf{M}_n(\Bbbk)$. Si existe una matriz inversible $C \in \mathsf{GL}_n(\Bbbk)$ tal que $B = C^{-1}AC$, entonces $\chi_B = \chi_A$.*

Demostración. En efecto, si existe una tal matriz, tenemos en $\mathsf{M}_n(\Bbbk(X))$ que

$$XI_n - B = XC^{-1}I_nC - C^{-1}AC = C^{-1}(XI_n - A)C$$

de manera que, según el Corolario 4.3.4, es

$$\chi_B = \det(XI_n - B) = \det C^{-1}(XI_n - A)C = \det(XI_n - A) = \chi_A,$$

como afirma la proposición. $\square$

5.3.5. Usando este resultado podemos extender la noción de polinomio característico a endomorfismos: si $f : V \to V$ es un endomorfismo de un espacio vectorial V de dimensión finita n y $\mathscr{B}$ es una base ordenada de V, llamamos ***polinomio característico de f*** y escribimos χ_f al polinomio característico de la matriz $[f]^{\mathscr{B}}_{\mathscr{B}}$, de manera que

$$\chi_f = \det(XI_n - [f]^{\mathscr{B}}_{\mathscr{B}}).$$

Por supuesto, para que esta definición tenga sentido el segundo miembro de esta igualdad tiene que depender solamente de f y no de la elección de la base $\mathscr{B}$. Para ver que esto es así, sea $\mathscr{B}'$ otra base ordenada de V y sea $C = C(\mathscr{B}', \mathscr{B})$ la matriz de cambio de base de $\mathscr{B}'$ a $\mathscr{B}$. Sabemos que

$$[f]^{\mathscr{B}'}_{\mathscr{B}'} = C(\mathscr{B}, \mathscr{B}') \cdot [f]^{\mathscr{B}}_{\mathscr{B}} \cdot C(\mathscr{B}', \mathscr{B}) = C^{-1} \cdot [f]^{\mathscr{B}}_{\mathscr{B}} \cdot C$$

y, por lo tanto, que las matrices $[f]^{\mathscr{B}}_{\mathscr{B}}$ y $[f]^{\mathscr{B}'}_{\mathscr{B}'}$ son conjugadas: la Proposición 5.3.4 nos dice entonces que estas dos matrices tienen el mismo polinomio característico, que es lo que queremos.

Una consecuencia inmediata de las definiciones y del Ejemplo 2.8.2(a) es que si $n \in \mathbb{N}$ y $A \in \mathsf{M}_n(\Bbbk)$, entonces el polinomio característico χ_A de A coincide con el polinomio característico χ_{f_A} del endomorfismo $f_A : x \in \Bbbk^n \mapsto Ax \in \Bbbk^n$ de $\Bbbk^n$ asociado a A.

5.3.6. Ejemplos.

(a) Sea V un espacio vectorial de dimensión finita y sea $f : V \to V$ un endomorfismo de V. Si $V_1, \dots, V_r$ son subespacios de V que son f-invariantes y tales que $V = V_1 \oplus \cdots \oplus V_r$, entonces

$$\chi_f = \chi_{f_{V_1}} \cdots \chi_{f_{V_r}}.$$

Para verlo, para cada $i \in [\![r]\!]$ sea $m_i = \dim V : i$ y sea $\mathscr{B}_i = (x_{1,1}, \dots, x_{1,m_i})$ una base ordenada de V_i. Entonces $\mathscr{B} = (x_{1,1}, \dots, x_{1,m_1}, \dots, x_{r,1}, \dots, x_{r,m_r})$ la base ordenada de V que se obtiene concantenando las bases $\mathscr{B}_1, \dots, \mathscr{B}_r$. La matriz de f con respecto a $\mathscr{B}$ es la matriz diagonal por bloques

$$[f]_{\mathscr{B}}^{\mathscr{B}} = \begin{pmatrix} [f_{V_1}]_{\mathscr{B}_1}^{\mathscr{B}_1} & & \\ & \ddots & \\ & & [f_{V_r}]_{\mathscr{B}_r}^{\mathscr{B}_r} \end{pmatrix},$$

así que el resultado es consecuencia de lo que vimos en el Ejemplo 5.3.3(b).

(b) De manera similar, si $f : V \to V$ es un endomorfismo de un espacio vectorial de dimensión finita y W es un subespacio f-invariante de V, entonces se tiene que

$$\chi_f = \chi_{f_W} \cdot \chi_{f^W}. \tag{6}$$

Si n y m son las dimensiones de V y de W, respectivamente, y $\mathscr{B} = (x_1, \dots, x_n)$ es una base ordenada de V tal que $\mathscr{B}' = (x_1, \dots, x_m)$ es una base ordenada de W, entonces sabemos que $\mathscr{B}'' = ([x_{m+1}], \dots, [x_n])$ es una base ordenada de V/W y que hay una matriz $C \in \mathsf{M}_{r,n}(\Bbbk)$ tal que la matriz de f con respecto a $\mathscr{B}$ es triangular superior por bloques, de la forma

$$[f]_{\mathscr{B}}^{\mathscr{B}} = \begin{pmatrix} [f_W]_{\mathscr{B}_W}^{\mathscr{B}_W} & C \\ 0 & [f^W]_{\mathscr{B}^W}^{\mathscr{B}^W} \end{pmatrix}.$$

La igualdad (6) es consecuencia de esto y del Ejemplo 5.3.3(b).

(c) Si V es un espacio vectorial de dimensión finita y $f : V \to V$ es un endomorfismo de V, entonces el polinomio característico χ_f de f coincide con el polinomio característico $\chi_{f^{\mathrm{t}}}$ de la función transpuesta $f^{\mathrm{t}} : V^* \to V^*$, esto es,

$$\chi_{f^{\mathrm{t}}} = \chi_f.$$

En efecto, si $\mathscr{B}$ es una base ordenada de V y $\mathscr{B}^*$ es la base de V^* dual a $\mathscr{B}$, entonces la Proposición 3.3.6 nos dice que la matriz $[f^{\mathrm{t}}]_{\mathscr{B}^*}^{\mathscr{B}^*}$ es la transpuesta de $[f]_{\mathscr{B}}^{\mathscr{B}}$ y el Ejemplo 5.3.3(d) nos dice que, en consecuencia, $[f^{\mathrm{t}}]_{\mathscr{B}^*}^{\mathscr{B}^*}$ y $[f]_{\mathscr{B}}^{\mathscr{B}}$ tienen el mismo polinomio característico.

(d) Sea V un espacio vectorial de dimensión finita, sea $f : V \to V$ un endomorfismo de V y supongamos que W es un subespacio de V que es f-invariante, de manera que podemos considerar la restricción $f_W : W \to W$ de f a W y la función lineal inducida $f^W : V/W \to V/W$ por f en el espacio cociente V/W. Afirmamos que el polinomio característico de f es el producto de los de f_W y de f^W,

$$\chi_f = \chi_{f_W} \cdot \chi_{f^W}.$$

Para verlo, **Hacer.**

$\diamond$

5.3.7. Nuestro interés en el polinomio característico de un endomorfismo reside en que vale la siguiente proposición:

Proposición. *Sea V un espacio vectorial de dimensión finita n y sea $f : V \to V$ un endomorfismo de V. El polinomio característico χ_f es mónico y tiene grado n. Si $\chi_f = X^n + c_1 X^{n-1} + \cdots + c_{n-1}X + c_n$ con $c_1, \dots, c_n \in \Bbbk$, entonces*

$$c_1 = -\operatorname{tr}(f), \qquad\qquad c_n = (-1)^n \det f.$$

Un escalar $\lambda \in \Bbbk$ es un autovalor de f si y solamente si es raíz de χ_f, y la multiplicidad geométrica de λ como autovalor de f es a lo sumo igual a la multiplicidad de λ como raíz de χ_f.

Demostración. Sea $\mathscr{B}$ una base ordenada de V y sea $(a_{i,j})$ la matriz $[f]^{\mathscr{B}}_{\mathscr{B}}$ de f con respecto a $\mathscr{B}$. Si denotamos $(b_{i,j})$ a la matriz $XI_n - [f]^{\mathscr{B}}_{\mathscr{B}}$, cuyas entradas son polinomios de $\Bbbk[X]$, la fórmula de Leibniz 4.4.14 nos dice que

$$\chi_f = \det(XI_n - [f]^{\mathscr{B}}_{\mathscr{B}}) = \sum_{\sigma \in S_n} \operatorname{sgn}(\sigma) b_{\sigma(1),1} \cdots b_{\sigma(n),n}. \tag{7}$$

Ahora bien, para cada $i, j \in [\![n]\!]$ es

$$b_{i,j} = \begin{cases} X - a_{i,j}, & \text{si } i = j; \\ -a_{i,j}, & \text{en caso contrario} \end{cases}$$

Así, el polinomio $b_{i,j}$ tiene grado 1 si $i = j$ y es o nulo o de grado 0 si $i \neq j$. Esto implica que para cada $\sigma \in S_n$ el término

$$\operatorname{sgn}(\sigma) b_{\sigma(1),1} \cdots b_{\sigma(n),n} \tag{8}$$

de la suma (7) correspondiente a σ es un polinomio que, si no es nulo, tiene de grado igual a la cantidad de elementos del conjunto $F_\sigma = \{k \in [\![n]\!] : \sigma(k) = k\}$ y es fácil ver que este conjunto F_σ

- tiene a lo sumo n elementos,
- tiene n elementos si y solamente si σ es la permutación identidad id_n de $[\![n]\!]$, y
- si tiene menos que n elementos entonces tiene a lo sumo $n-2$.

Escribiendo la suma (7) en la forma

$$\chi_f = b_{1,1} \cdots b_{n,n} + \sum_{\substack{\sigma \in S_n \\ \sigma \neq \mathrm{id}_n}} \operatorname{sgn}(\sigma) b_{\sigma(1),1} \cdots b_{\sigma(n),n},$$

resulta entonces que el segundo sumando en el miembro izquierdo de la igualdad es o nulo o que tiene grado estrictamente menor que $n-1$, mientras que el primero tiene grado exactamente igual a n. Esto implica, por supuesto, que el grado de χ_f es exactamente n y, más aún, que los coeficientes de X^n y de X^{n-1} en χ_f coinciden con los de X^n y de X^{n-1} en

$$b_{1,1} \cdots b_{n,n} = (X - a_{1,1}) \cdots (X - a_{n,n}),$$

que son 1 y $-(a_{1,1} + \cdots + a_{n,n})$, respectivamente: esto nos dice que el polinomio χ_f es mónico y que el coeficiente con el que aparece X^{n-1} en él es precisamente $-\operatorname{tr}(f)$.

Por otro lado, para cada $\sigma \in S_n$ el término constante del sumando (8) de la suma (7) es $(-1)^n \operatorname{sgn}(\sigma) a_{\sigma(1),1} \cdots a_{\sigma(n),n}$, así que el término constante del polinomio χ_f es

$$(-1)^n \sum_{\sigma \in S_n} \operatorname{sgn}(\sigma) a_{\sigma(1),1} \cdots a_{\sigma(n),n} = (-1)^n \det[f]_{\mathscr{B}}^{\mathscr{B}} = (-1)^n \det f.$$

Sea ahora $\lambda \in \Bbbk$ un escalar. Como

$$\chi_f(\lambda) = \det(\lambda I_n - [f]_{\mathscr{B}}^{\mathscr{B}}) = \det[\lambda \mathrm{id}_V - f]_{\mathscr{B}}^{\mathscr{B}} = \det(\lambda \mathrm{id}_V - f),$$

se sigue inmediatamente del Proposición 5.3.1 que λ es un autovalor de f si y solamente si $\chi_f(\lambda) = 0$.

Sea r la multiplicidad geométrica de λ como autovalor de f, de manera que $r = \dim E_\lambda(f)$. Sin perdida de generalidad, podemos suponer que la base ordenada $\mathscr{B} = (x_1, \ldots, x_n)$ de V es tal que $(x_1, \ldots, x_r)$ es una base de $E_\lambda(f)$ y entonces, como para cada $i \in [\![r]\!]$ es $f(x_i) = \lambda x_i$, la matriz de f con respecto a la base $\mathscr{B}$ tiene una descomposición en bloques de la forma

$$[f]_{\mathscr{B}}^{\mathscr{B}} = \begin{pmatrix} \lambda I_r & B \\ 0 & C \end{pmatrix}$$

con $B \in \mathrm{M}_{r,n-r}(\Bbbk)$ y $C \in \mathrm{M}_r(\Bbbk)$. Se sigue de esto que

$$\chi_f = \det(XI_n - [f]_{\mathscr{B}}^{\mathscr{B}}) = \det \begin{pmatrix} (X-\lambda)I_r & -B \\ 0 & XI_{n-r} - C \end{pmatrix} = (X-\lambda)^r \chi_C$$

y, en consecuencia, que la multiplicidad de λ como raíz de χ_f es por lo menos r. Esto prueba la última afirmación de la parte (*i*) de la proposición. □

5.3.8. Llamamos ***multiplicidad algebraica*** de un escalar $\lambda \in \Bbbk$ como autovalor de un endomorfismo f de un espacio vectorial de dimensión finita o de una matriz A a la multiplicidad con que λ es raíz del polinomio característico χ_f o χ_A. La Proposición 5.3.7 nos dice, usando este lenguaje, que la multiplicidad geométrica de un autovalor es menor o igual a su multiplicidad algebraica.

5.3.9. De la Proposición 5.3.7 obtenemos fácilmente una condición suficiente para la diagonalizabilidad de un endomorfismo o una matriz:

Proposición. *Sea V un espacio vectorial de dimensión finita n y sea $f : V \to V$ un endomorfismo de V. Si el polinomio característico de f tiene n raíces simples en $\Bbbk$, entonces f es diagonalizable.*

La condición que da esta proposición no es necesaria: por ejemplo, la función identidad $\mathrm{id}_{\Bbbk^n} : \Bbbk^n \to \Bbbk^n$ es diagonalizable pero su polinomio característico, $(X-1)^n$, no tiene raíces simples salvo si $n = 1$.

Demostración. Supongamos que el polinomio característico χ_f tiene n raíces simples en $\Bbbk$. De acuerdo a la Proposición 5.3.7, cada una de esas raíces es un autovalor de f y entonces f posee n autovectores de autovalores distintos dos a dos. El conjunto de estos autovectores es, según la Proposición 5.1.4, linealmente independiente y se trata, por lo tanto, de una base de V: esto significa que f es diagonalizable. □

5.3.10. Ejemplos.

(a) Si $p \in \Bbbk[X]$ es un polinomio mónico de grado n que tiene todas sus raíces en $\Bbbk$ y simples, entonces la matriz compañera $C(p)$ es diagonalizable: en efecto, el polinomio característico de $C(p)$ es precisamente p, así que esto sigue de la Proposición 5.3.9. Observemos que ya habíamos obtenido esta conclusión en el Ejemplo 5.2.3.

(b) Fijemos ahora $n \in \mathbb{N}$ y consideremos la matriz $A_n = (a_{i,j}) \in \mathsf{M}_n(\Bbbk)$ que para cada $i, j \in [\![n]\!]$ tiene entrada (i, j)-ésima dada por

$$a_{i,j} = \begin{cases} -1, & \text{si } |i-j| = 1; \\ 0, & \text{en caso contrario.} \end{cases}$$

Por ejemplo, cuando n es 5 la matriz es

$$A_5 = \begin{pmatrix} 0 & -1 & 0 & 0 & 0 \\ -1 & 0 & -1 & 0 & 0 \\ 0 & -1 & 0 & -1 & 0 \\ 0 & 0 & -1 & 0 & -1 \\ 0 & 0 & 0 & -1 & 0 \end{pmatrix}.$$

Si $n > 2$, para calcular el polinomio característico de A_n podemos considerar el desarrollo de Laplace a lo largo de la primera de las filas de la matriz $XI_n - A_n \in \mathsf{M}_n(\Bbbk(X))$,

$$\chi_{A_n} = \begin{vmatrix} X & 1 & 0 & 0 & \cdots \\ 1 & X & 1 & 0 & \cdots \\ 0 & 1 & X & 1 & \cdots \\ 0 & 0 & 1 & X & \cdots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{vmatrix} = X \begin{vmatrix} X & 1 & 0 & \cdots \\ 1 & X & 1 & \cdots \\ 0 & 1 & X & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{vmatrix} - \begin{vmatrix} 1 & 1 & 0 & \cdots \\ 0 & X & 1 & \cdots \\ 0 & 1 & X & \cdots \\ \vdots & \vdots & \vdots & \ddots \end{vmatrix}.$$

El primero de estos dos determinantes es precisamente el que calcula el polinomio característico de la matriz A_{n-1}; por otro lado, desarrollando el segundo a lo largo de su primera columna vemos que es igual al polinomio característico de A_{n-2}. Concluimos así que

$$\chi_{A_n} = X\chi_{A_{n-1}} - \chi_{A_{n-2}} \qquad \text{si } n > 2. \tag{9}$$

Por otro lado, calculando directamente vemos que

$$\chi_{A_1} = X, \qquad \chi_{A_2} = X^2 - 1, \tag{10}$$

Estas dos igualdades y la relación de recurrencia (9) nos permiten calcular los polinomios χ_{A_n} recursivamente. En el Cuadro 5.1 de la página siguiente listamos los primeros.

Para cada $n \in \mathbb{N}$ consideremos el polinomio U_n tal que

$$U_n(X) = \chi_{A_n}(2X).$$

Como consecuencia de (9) y de (10) se tiene que

$$U_1 = 2X, \qquad U_2 = 4X^2 - 1, \qquad U_n = 2XU_{n-1} - U_{n-1} \quad \text{si } n > 2.$$

n	χ_{A_n}	U_n
1	X	$2X$
2	$X^2 - 1$	$4X^2 - 1$
3	$X^3 - 2X$	$8X^3 - 4X$
4	$X^4 - 3X^2 + 1$	$16X^4 - 12X^2 + 1$
5	$X^5 - 4X^3 + 3X$	$32X^5 - 32X^3 + 6X$
6	$X^6 - 5X^4 + 6X^2 - 1$	$64X^6 - 80X^4 + 24X^2 - 1$
7	$X^7 - 6X^5 + 10X^3 - 4X$	$128X^7 - 192X^5 + 80X^3 - 8X$
8	$X^8 - 7X^6 + 15X^4 - 10X^2 + 1$	$256X^8 - 448X^6 + 240X^4 - 40X^2 + 1$
9	$X^9 - 8X^7 + 21X^5 - 20X^3 + 5X$	$512X^9 - 1024X^7 + 672X^5 - 160X^3 + 10X$

Cuadro 5.1. Los polinomios característicos de las matrices A_n del ejemplo 5.3.10(b) y los polinomios de Chebyshev de segunda especie.

El polinomio U_n se llama el n-ésimo ***polinomio de Chebyshev de segunda especie***, por Pafnuti Chebyshev (1821–1894, Rusia); es usual iniciar la sucesión de estos polinomios con $U_0 = 1$.

Supongamos desde ahora que el cuerpo de base con el que estamos trabajando es $\Bbbk = \mathbb{R}$. Una inducción evidente muestra que para cada $\theta \in \mathbb{R} \setminus \pi\mathbb{Z}$ es

$$U_n(\cos\theta) = \frac{\sin(n+1)\theta}{\sin\theta}.$$

Una consecuencia inmediata de esto es que el polinomio χ_{A_n} se anula en los n números reales

$$2\cos\frac{k\pi}{n+1} \qquad \text{con } k \in [\![n]\!],$$

que son, por lo tanto, sus n raíces. Estos n números son distintos dos a dos, así que el polinomio característico χ_{A_n} tiene sus n raíces en $\mathbb{R}$ y éstas son todas simples. La Proposición 5.3.9 nos dice entonces que la matriz A_n es diagonalizable.

(c) Fijemos otra vez un entero positivo $n \in \mathbb{N}$ y trabajemos sobre el cuerpo $\mathbb{R}$ de los números reales. Si p es un elemento del espacio vectorial $\mathbb{R}[X]_{\leq n}$, entonces $xp' - p''$ también lo es y podemos por lo tanto considerar la función lineal

$$H : p \in \mathbb{R}[X]_{\leq n} \mapsto xp' - p'' \in \mathbb{R}[X]_{\leq n}.$$

Calculando, vemos inmediatamente que para cada $i \in [\![0, n]\!]$ es

$$H(X^i) = iX^i - i(i-1)X^{i-2},$$

así que si la matriz de H con respecto a la base ordenada $\mathscr{B} = (1, X, \dots, X^n)$ de $\mathbb{R}[X]_{\leq n}$ es la

n	He_n
0	1
1	X
2	$X^2 - 1$
3	$X^3 - 3X$
4	$X^4 - 6X^2 + 3$
5	$X^5 - 10X^3 + 15X$
6	$X^6 - 15X^4 + 45X^2 - 15$
7	$X^7 - 21X^5 + 105X^3 - 105X$
8	$X^8 - 28X^6 + 210X^4 - 420X^2 + 105$
9	$X^9 - 36X^7 + 378X^5 - 1260X^3 + 945X$
10	$X^{10} - 45X^8 + 630X^6 - 3150X^4 + 4725X^2 - 945$

Cuadro 5.2. Los primeros diez polinomios de Hermite.

matriz triangular superior

$$[H]_{\mathscr{B}}^{\mathscr{B}} = \begin{pmatrix} 0 & 0 & -2 & 0 & \cdots & 0 \\ & 1 & 0 & -6 & \ddots & \vdots \\ & & 2 & 0 & \ddots & 0 \\ & & & 3 & \ddots & -n(n-1) \\ & & & & \ddots & 0 \\ & & & & & n \end{pmatrix}.$$

De acuerdo al Ejemplo 5.3.3(a), entonces, los autovalores de f son los números $0, 1, \ldots, n$. Estos autovalores son $n+1$ en número y son distintos dos a dos, así que, como $\dim \mathbb{R}[X]_{\leq n} = n+1$, la Proposición 5.3.9 nos dice que la función H es diagonalizable.

Sea $k \in [\![0, n]\!]$. La multiplicidad algebraica de k como autovalor de H es 1, así que la multiplicidad geométrica de k es exactamente 1: existe entonces un único autovector mónico $p \in \Bbbk[X]$ de H de autovalor k. Si d es el grado de p, entonces $p = X^d + \sum_{i=1}^{d-1} c_i X^i$ para ciertos escalares $c_1, \ldots, c_{d-1} \in \mathbb{R}$ y

$$kX^d + k\sum_{i=1}^{d-1} c_i X^i i = kp = H(p) = dX^d \underbrace{- d(d-1)X^{d-2} + H\left(\sum_{i=1}^{d-1} c_i X^i\right)}.$$

La suma de los términos marcados en el último miembro de esta cadena de igualdades es un polinomio o nulo o de grado menor que d: el coeficiente de X^d en este último miembro es entonces d y, comparando con el primer miembro de la igualdad, vemos que tiene que ser

$d = k$. Así, hemos probado que

> *para cada $k \in [\![0, n]\!]$ la función lineal H tiene exactamente un autovector mónico de autovalor k y grado k.*

Llamamos a ese autovector el ***k-ésimo polinomio de Hermite***, por Charles Hermite (1822–1901, Francia), y lo escribimos He_k. De acuerdo a lo anterior, se trata de la única solución de la ecuación diferencial ordinaria

$$p'' - Xp' + kp = 0$$

que es un polinomio mónico. En el Cuadro 5.2 están tabulados los diez primeros. ◇

5.3.11. La consecuencia más importante de la Proposición 5.3.7 es que nos permite probar, bajo hipótesis razonables sobre el cuerpo $\Bbbk$, que un endomorfismo de un espacio vectorial o una matriz posee autovalores y autovalores.

Proposición. *Supongamos que el cuerpo $\Bbbk$ es algebraicamente cerrado. Si $f : V \to V$ es un endomorfismo de un espacio vectorial V de dimensión finita, entonces f posee al menos un autovalor.*

Demostración. Como el cuerpo $\Bbbk$ es algebraicamente cerrado, el polinomio característico χ_f, que es un elemento de $\Bbbk[X]$ posee una raíz en $\Bbbk$, esto es, existe $\lambda \in \Bbbk$ tal que $\chi_f(\lambda) = 0$, y de acuerdo a la Proposición 5.3.7, esto implica que λ es un autovalor de f. □

5.3.12. Que el cuerpo sea algebraicamente cerrado es, de hecho, una condición necesaria para que valga la conclusión de la Proposición 5.3.11 con total generalidad. En efecto, si $\Bbbk$ no es algebraicamente cerrado, entonces existe un polinomio $p \in \Bbbk[X]$ que no tiene ninguna raíz en $\Bbbk$ y la Proposición 5.3.7 nos dice que la matriz compañera $C(p)$ no tiene ningún autovalor, ya que, de acuerdo al ejemplo 5.3.3(c), su polinomio característico es precisamente p. Por ejemplo, podemos tomar $p = X^2 + 1$ si $\Bbbk$ es $\mathbb{Q}$ o $\mathbb{R}$, o $p = X^2 + X + 1$ si $\Bbbk = \mathbb{F}_2$.

5.3.13. Una aplicación de la Proposición 5.3.11 es:

Proposición. *Supongamos que el cuerpo $\Bbbk$ es algebraicamente cerrado. Sea V un espacio vectorial de dimensión finita y sea $f : V \to V$ un endomorfismo de V. Existe una base ordenada $\mathscr{B}$ de V tal que la matriz $[f]^{\mathscr{B}}_{\mathscr{B}}$ de f con respecto a $\mathscr{B}$ es triangular superior.*

Demostración. Procedemos haciendo inducción con respecto a la dimensión de V, notando que cuando $\dim V \leq 1$ no hay nada que probar.

Supongamos que V es un espacio vectorial de dimensión $n > 1$. La función transpuesta $f^{\mathrm{t}} : V^* \to V^*$ es un endomorfismo de un espacio vectorial de dimensión finita, así que, como el cuerpo $\Bbbk$ es algebraicamente cerrado, la Proposición 5.3.11 nos dice que f^{t} posee un autovalor $\lambda \in \Bbbk$ y, por lo tanto, que existe un autovector $\phi \in V^*$ de f^t de autovalor λ, de manera que $f^{\mathrm{t}}(\phi) = \lambda\phi$.

La función lineal $\phi : V \to \Bbbk$ no es nula, porque es un autovector de f^{t}, así que es sobreyectiva y, en consecuencia, su núcleo $V' = \mathsf{Nu}(\phi)$ es un subespacio de V de dimensión $n - 1$. Si $x \in V'$, entonces

$$\phi(f(v)) = f^{\mathrm{t}}(\phi)(v) = (\lambda\phi)(v) = \lambda\phi(v) = 0,$$

de manera que $f(v)$ también es un elemento de V'. Esto significa que la función f se restringe a un endomorfismo $f_{V'} : V' \to V'$ de V'. Como $\dim V' = n-1$, podemos suponer inductivamente que existe una base ordenada $\mathscr{B}' = (x_1, \dots, x_{n-1})$ de V' tal que la matriz $[f_{V'}]^{\mathscr{B}'}_{\mathscr{B}'}$ es triangular superior. Por otro lado, sabemos que podemos completar $\mathscr{B}'$ a una base ordenada de V, esto es, que existe un vector $x_n \in V$ tal que $\mathscr{B} = (x_1, \dots, x_{n-1}, x_n)$ es una base ordenada de V y, en particular, existen escalares $b_1, \dots, b_n \in \Bbbk$ tales que $f(x_n) = b_1x_1 + \cdots + b_nx_n$. Ahora bien, si consideramos la matriz

$$B = \begin{pmatrix} b_1 \\ \vdots \\ b_{n-1} \end{pmatrix} \in \mathsf{M}_{n-1,1}(\Bbbk),$$

entonces la matriz de f con respecto a la base $\mathscr{B}$ tiene una descomposición en bloques de la forma

$$[f]^{\mathscr{B}}_{\mathscr{B}} = \begin{pmatrix} [f_{V'}]^{\mathscr{B}'}_{\mathscr{B}'} & B \\ 0 & b_n \end{pmatrix}.$$

Como $[f_{V'}]^{\mathscr{B}'}_{\mathscr{B}'}$ es triangular superior, la matriz $[f]^{\mathscr{B}}_{\mathscr{B}}$ es también triangular superior y esto prueba lo que queremos. □

§4. Homomorfismos de álgebras e ideales

5.4.1. Si A y B son dos álgebras, una función lineal $f : A \to B$ es un ***homomorfismo de álgebras*** si $f(1_A) = 1_B$ y cada vez x e y son elementos de A se tiene que

$$f(x \cdot y) = f(x) \cdot f(y).$$

En ese caso es fácil ver que para cada $x \in A$ se tiene que $f(x^i) = f(x)^i$ para todo $i \in \mathbb{N}_0$.

5.4.2. Proposición. *Sea A un álgebra. Si $a \in A$, entonces existe exactamente un homomorfismo de álgebras $\varepsilon_a : \Bbbk[X] \to A$ tal que $\varepsilon_a(X) = a$.*

Demostración. Como el conjunto $\{X^i : i \in \mathbb{N}_0\}$ es una base de $\Bbbk[X]$, sabemos que existe una función lineal $\varepsilon_a : \Bbbk[X] \to A$ tal que $\varepsilon_a(X^i) = a^i$ para cada $i \in \mathbb{N}_0$. Mostremos que ε_a es un homomorfismo de álgebras —esto probará la afirmación de existencia.

Que $\varepsilon_a(1_{\Bbbk[X]}) = 1_A$ es consecuencia de la definición de ε_a. Sean p y q dos elementos de $\Bbbk[X]$. Existen $m, n \in \mathbb{N}_0$ y escalares $\lambda_0, \dots, \lambda_m$ y $\mu_0, \dots, \mu_n$ en $\Bbbk$ tales que $p = \sum_{i=0}^m \lambda_i X^i$ y $q = \sum_{j=1}^n \mu_j X^j$, y entonces

$$p \cdot q = \sum_{i=0}^m \lambda_i X^i \cdot \sum_{j=0}^n \mu_j X^j = \sum_{k=0}^{m+n} \left(\sum_{\substack{0 \le i \le m \\ 0 \le j \le n \\ i+j=k}} \lambda_i \mu_j \right) X^k. \tag{11}$$

Por la forma en que fue definida la función ε_a sabemos que $\varepsilon_a(p) = \sum_{i=0}^m \lambda_i a^i$ y $\varepsilon_a(q) = \sum_{j=0}^n \mu_j a^j$, y se sigue de esto que

$$\varepsilon_a(p) \cdot \varepsilon_a(q) = \sum_{i=0}^m \lambda_i a^i \cdot \sum_{j=0}^n \mu_j a^j = \sum_{k=0}^{m+n} \left(\sum_{\substack{0 \le i \le m \\ 0 \le j \le n \\ i+j=k}} \lambda_i \mu_j \right) a^k.$$

De manera similar, de la igualdad (11) vemos

$$\varepsilon_a(p \cdot q) = \sum_{k=0}^{m+n} \left(\sum_{\substack{0 \le i \le m \\ 0 \le j \le n \\ i+j=k}} \lambda_i \mu_j \right) a^k$$

y comparando esta expresión con la que obtuvimos para $\varepsilon_a(p) \cdot \varepsilon_a(q)$ es claro que estos dos elementos de A son iguales.

Supongamos ahora, para probar la unicidad, que $\varepsilon, \eta : \Bbbk[X] \to A$ son dos homomorfismos de álgebras y que $\varepsilon(X) = a = \eta(X)$. Es entonces $\varepsilon(1_{\Bbbk[X]}) = 1_A = \eta(1_{\Bbbk[X]})$ y, para cada $i \in \mathbb{N}$,

$$\varepsilon(X^i) = \varepsilon(X)^i = a^i = \eta(X)^i = \eta(X^i).$$

Vemos así que ε y η coinciden sobre todos los elementos de la base $\{X^i : i \in \mathbb{N}_0\}$ de $\Bbbk[X]$ y, por lo tanto, que son iguales. $\square$

5.4.3. Cada vez que A es un álgebra asociativa y a un elemento de A, la Proposición 5.4.2 nos provee un homomorfismo de álgebras $\varepsilon_a : \Bbbk[X] \to A$ completamente determinado por la condición de que sea $\varepsilon_a(X) = a$. Si $p \in \Bbbk[X]$ escribimos siempre $p(a)$ en lugar de $\varepsilon_a(p)$ y decimos que $p(a)$ se obtiene ***evaluando el polinomio*** p ***en*** a.

5.4.4. En general, la multiplicación de un álgebra no es conmutativa. Vale sin embargo la siguiente observación importante, que usaremos muchas veces:

Lema. *Sean A y B dos álgebras y sea $f : A \to B$ un homomorfismo de álgebras. Si A es conmutativa y b y b' son elementos de $f(A)$, entonces $bb' = b'b$.*

Demostración. Como b y b' son elementos de $f(A)$, existe a y a' en A tales que $b = f(a)$ y $b' = f(a')$, y entonces

$$bb' = f(a)f(a') = f(aa') = f(a'a) = f(a')f(a) = b'b,$$

porque A es conmutativa. $\square$

5.4.5. Si A es un álgebra, un subespacio I de A es un ***ideal*** si cada vez que $x \in I$ e $y \in A$ se tiene que xy e yx está en I. Es fácil exhibir ejemplos de ideales:

Proposición. *Sean A y B dos álgebras. Si $f : A \to B$ es un homomorfismo de álgebras, entonces el núcleo $\mathsf{Nu}(f)$ es un ideal de A.*

Demostración. Si $x \in \mathsf{Nu}(f)$ e $y \in A$, entonces

$$f(xy) = f(x)f(y) = 0f(y) = 0,$$

porque f es un homomorfismo y, de manera similar, $f(yx) = 0$. Como $\mathsf{Nu}(f)$ es un subespacio de A, esto nos dice que se trata de un ideal. □

5.4.6. En general, un álgebra tiene muchos ideales. El siguiente resultado nos describe todos los de un álgebra de polinomios:

Proposición. *Si I es un ideal no nulo de $\Bbbk[X]$, entonces existe un único polinomio mónico $m \in I$ tal que para cada $p \in \Bbbk[X]$ se tiene que*

$$p \textit{ pertenece a } I \textit{ si y solamente si } p \textit{ es divisible por } m \tag{12}$$

y, de hecho, es $I = \{pm : p \in \Bbbk[X]\}$.

Demostración. Como hay elementos no nulos en I, tiene sentido considerar el número

$$d = \min\{\mathsf{gr}(p) : p \in I \setminus 0\}$$

y existe en I un elemento m no nulo tal que $\operatorname{gr} m = d$; a menos de cambiar a m por uno de sus múltiplos escalares, podemos suponer que m es mónico.

Sea $p \in \Bbbk[X]$ un polinomio. Si p es divisible por m, existe $q \in \Bbbk[X]$ tal que $p = mq$ y entonces p pertenece a I porque éste es un ideal de $\Bbbk[X]$. Para ver la recíproca, supongamos ahora que p pertenece a I. Como m no es nulo y tiene grado d, sabemos que existen polinomios s y r en $\Bbbk[X]$ tales que $p = sm + r$ y o bien $r = 0$ o bien $\mathsf{gr}(r) < d$. En particular, tenemos que $r = p - sm$ y, como I es un ideal, se sigue de esto que $r \in I$. Esto implica —en vista de la forma en que elegimos el número d— que no puede ser que r no sea nulo, porque su grado en ese caso es menor que d. Así, es $r = 0$, $p = sm$ y, en definitiva, m divide a p.

Para ver la unicidad de m, supongamos que $\tilde{m}$ es otro elemento mónico de $\Bbbk[X]$ que tiene la propiedad descripta en el enunciado. Como m y $\tilde{m}$ tienen ambos esta propiedad y pertenecen a I, se dividen mutuamente y esto y el hecho de que ambos polinomios son mónicos implica, por supuesto, que $\tilde{m} = m$.

Como I es un ideal y $m \in I$, es claro que I contiene a $\{pm : p \in \Bbbk[X]\}$ y la contención recíproca es consecuencia inmediata de que m posee la propiedad (12) del enunciado. Esto muestra que vale la última afirmación de la proposición. □

§5. Polinomios y endomorfismos

5.5.1. Proposición. *Sea V un espacio vectorial, sea $f : V \to V$ un endomorfismo de V y sea $p \in \Bbbk[X]$. Si W es un subespacio f-invariante de V, entonces W es $p(f)$-invariante, de manera que podemos considerar su restricción $p(f)_W : W \to W$ a W, y vale que $p(f_W) = p(f)_W$.*

Demostración. Supongamos que $p = a_d X^d + \cdots + a_0$, con $d \in \mathbb{N}_0$ y $a_d, \ldots, a_0 \in \Bbbk$ y sea W un subespacio f-invariante de V. Si $x \in W$ e $i \in \mathbb{N}_0$, entonces claramente $f^i(x) \in W$ y se sigue de esto que para cada $x \in W$ es

$$p(f)(x) = (a_d f^d + \cdots + a_1 d + a_0 \mathrm{id}_V)(x) = a_d f^d(x) + \cdots + a_1 f(x) + a_0 x \in W,$$

esto es, que W es $p(f)$-invariante. Si $p(f)_W : W \to W$ es la restricción de $p(f)$ a W y $x \in W$, entonces para cada $i \in \mathbb{N}_0$ es $(f_W)^i(x) = f^i(x)$, asi que

$$p(f_W)(x) = a_d f^d(x) + \cdots + a_1 f(x) + a_0 x = p(f)(x) = p(f)_W(x),$$

de manera que $p(f_W) = p(f)_W$, como afirma la proposición. □

5.5.2. Proposición. *Si V es un espacio vectorial y $f : V \to V$ es un endomorfismo, cada vez que p y q son dos polinomios de $\Bbbk[X]$ se tiene que $p(f) \circ q(f) = q(f) \circ p(f)$.*

Demostración. Sea $\varepsilon_f : \Bbbk[X] \to \mathrm{End}(V)$ el homomorfismo de álgebras tal que $\varepsilon_f(X) = X$. La afirmación es consecuencia inmediata del Lema 5.4.4, ya que por definición $p(f) = \varepsilon_f(p)$ y $q(f) = \varepsilon_f(q)$ están en la imagen de ε_f y el álgebra $\Bbbk[X]$ es conmutativa. □

5.5.3. Si $p \in \Bbbk[X]$ y $S \subseteq \Bbbk$ es un conjunto de escalares, escribimos $f(S)$ al conjunto $\{f(\lambda) : \lambda \in S\}$.

Proposición. *Sea V un espacio vectorial, sea $f : V \to V$ un endomorfismo de V y sea $p \in \Bbbk[X]$.*

(*i*) *Si x es un autovector de f de autovalor λ, entonces x es un autovector de $p(f)$ de autovalor $p(\lambda)$.*

(*ii*) *Se tiene que $\mathrm{Spec}(p(f)) \supseteq p(\mathrm{Spec}(f))$ y si el cuerpo $\Bbbk$ es algebraicamente cerrado y V tiene dimensión finita entonces vale, de hecho, la igualdad.*

Demostración. Si p es un polinomio constante, entonces $p(f)$ es un múltiplo escalar de id_V y todas las afirmaciones son inmediatas. Supongamos entonces que p no es constante, que $p = a_d X^d + \cdots + a_0$, con $a_0, \ldots, a_d \in \Bbbk$ y que $a_d \neq 0$, con lo que el grado de p es exactamente d.

(*i*) Si x es un autovector de f de autovector λ, entonces $f^i(x) = \lambda^i x$ para cada $i \in \mathbb{N}_0$ y

$$\begin{aligned} p(f)(x) &= (a_d f^d + \cdots + a_1 f + a_0 \mathrm{id}_V)(x) \\ &= a_d \lambda^d x + \cdots + a_1 \lambda x + a_0 x \\ &= (a_d \lambda^d + \cdots + a_1 \lambda + a_0) x \\ &= p(\lambda) x, \end{aligned}$$

de manera que x es un autovector de $p(f)$ de autovector $p(\lambda)$.

(*ii*) Que el conjunto $p(\mathsf{Spec}(f))$ está contenido en $\mathsf{Spec}(p(f))$ es consecuencia inmediata de la parte (*i*) de la proposición. Supongamos ahora que el cuerpo $\Bbbk$ es algebraicamente cerrado y que V tiene dimensión finita n, y sea $\mu \in \mathsf{Spec}(p(f))$. El polinomio $p - \mu$ se factoriza en $\Bbbk[X]$ como producto de factores lineales, de manera que existen $\alpha_0, \dots, \alpha_d \in \Bbbk$ tales que

$$p - \mu = \alpha_0(X - \alpha_1)\cdots(X - \alpha_d). \tag{13}$$

Observemos que como p no es constante, $\alpha_0 \neq 0$. Evaluando ambos lados de la igualdad en f vemos que

$$p(f) - \mu\mathsf{id}_V = \alpha_0(f - \alpha_1\mathsf{id}_V)\cdots(f - \alpha_d\mathsf{id}_V)$$

y, tomando ahora determinantes, que

$$\mathsf{det}(p(f) - \mu\mathsf{id}_V) = \alpha_0^n \,\mathsf{det}(f - \alpha_1\mathsf{id}_V)\cdots \mathsf{det}(f - \alpha_d\mathsf{id}_V).$$

Como μ es un autovalor de $p(f)$, el lado izquierdo de esta igualdad es nulo y entonces alguno de los factores que aparecen en el lado derecho tiene que ser nulo. Como $\alpha_0^n \neq 0$, vemos que existe $i \in [\![d]\!]$ tal que $\mathsf{det}(f - \alpha_i\mathsf{id}_V) = 0$, esto es, tal que α_i es un autovalor de f. Evaluando ahora ambos lados de la igualdad (13) en α_u, vemos que $p(\alpha_i) - \mu = 0$, así que $\mu = p(\alpha_i) \in \{p(\lambda) : \lambda \in \mathsf{Spec}(f)\}$. □

5.5.4. Proposición. *Sea V un espacio vectorial de dimensión finita.*

(*i*) *Un endomorfismo $f : V \to V$ es un automorfismo de V si y solamente si $0 \notin \mathsf{Spec}(f)$, y en ese caso es $\mathsf{Spec}(f^{-1}) = \{\lambda^{-1} : \lambda \in \mathsf{Spec}(f)\}$.*

(*ii*) *Si $f, g : V \to V$ son endomorfismos de V, entonces $\mathsf{Spec}(fg) = \mathsf{Spec}(gf)$.*

Demostración. (*i*) Sea $f : V \to V$ un endomorfismo de V. El escalar 0 pertenece a $\mathsf{Spec}(f)$ si y solamente si $\mathsf{det}(0\,\mathsf{id}_V - f) = \mathsf{det}(f) = 0$, y esto ocurre si y solamente si f no es un isomorfismo: esto prueba la primera afirmación.

Supongamos ahora que f es un isomorfismo y sea $\lambda \in \mathsf{Spec}(f)$, de manera que $\lambda\mathsf{id}_V - f$ no es inversible. Como f sí es inversible, tiene que ser $\lambda \neq 0$ y como

$$\lambda f(\lambda^{-1}\mathsf{id}_V - f^{-1}) = -(\lambda\mathsf{id}_V - f),$$

tomando determinantes vemos que

$$\mathsf{det}(\lambda f) \cdot \mathsf{det}(\lambda^{-1}\mathsf{id}_V - f^{-1}) = (-1)^{\mathsf{dim}\,V}\,\mathsf{det}(\lambda\mathsf{id}_V - f) = 0.$$

Es $\mathsf{det}(\lambda f) \neq 0$ porque λf es un automorfismo de V, y entonces tiene que ser $\mathsf{det}(\lambda^{-1}\mathsf{id}_V - f^{-1}) = 0$. Esto implica que $\lambda^{-1} \in \mathsf{Spec}(f^{-1})$ y prueba, en definitiva, que

$$\{\lambda^{-1} : \lambda \in \mathsf{Spec}(f)\} \subseteq \mathsf{Spec}(f^{-1}).$$

Reemplazando en esto f por f^{-1} y recordando que $(f^{-1})^{-1} = f$, vemos que también

$$\{\lambda^{-1} : \lambda \in \mathsf{Spec}(f^{-1})\} \subseteq \mathsf{Spec}(f).$$

De estas dos inclusiones es inmediato obtener la igualdad que aparece en el enunciado.

(ii) Sean $f, g : V \to V$ dos endomorfismos de V, sea $\lambda \in \Bbbk \setminus \operatorname{Spec}(fg)$ y mostremos que $\lambda \in \Bbbk \setminus \operatorname{Spec}(gf)$. Esto implicará, por supuesto, que $\operatorname{Spec}(gf) \subseteq \operatorname{Spec}(fg)$ y, por simetría, que ambos espectros son iguales. Consideramos dos casos:

- Supongamos primero que $\lambda = 0$. El endomorfismo fg es un automorfismo de V y, en consecuencia,

$$\det(gf) = \det g \cdot \det f = \det f \cdot \det g = \det(fg) \neq 0.$$

  Esto implica que gf también es un automorfismo de V y, por lo tanto, que $\lambda \in \Bbbk \setminus \operatorname{Spec}(gf)$.
- Supongamos, en segundo lugar, que $\lambda \neq 0$. Como λ no es un autovalor de fg, el endomorfismo $\lambda \mathrm{id}_V - fg$ es inversible y existe un endomorfismo $h : V \to V$ tal que

$$h \cdot (\lambda \mathrm{id}_V - fg) = (\lambda \mathrm{id}_V - fg) \cdot h = \mathrm{id}_V.$$

  Afirmamos que el endomorfismo $k = \lambda^{-1}(\mathrm{id}_V + ghf)$ es inverso de $\lambda \mathrm{id}_V - gf$. En efecto,

$$\begin{aligned} k \cdot (\lambda \mathrm{id}_V - gf) &= \lambda^{-1}(\mathrm{id}_V + ghf)(\lambda \mathrm{id}_V - gf) \\ &= \mathrm{id}_V - \lambda^{-1} gf + ghf - \lambda^{-1} ghfgf \\ &= \mathrm{id}_V - \lambda^{-1} gf + \lambda^{-1} g \cdot \underbrace{h \cdot (\lambda h - hfg)}\cdot f \end{aligned}$$

  y, como la expresión marcada es igual a id_V, esto es

$$= \mathrm{id}_V - \lambda^{-1} gf + \lambda^{-1} gf = \mathrm{id}_V,$$

  y un cálculo similar muestra que $(\lambda \mathrm{id}_V - gf) \cdot k = \mathrm{id}_V$. En particular, como $\lambda \mathrm{id}_V - gf$ es inversible, vemos que $\lambda \in \Bbbk \setminus \operatorname{Spec}(gf)$.

Así, en cualquier caso vemos que $\lambda \in \Bbbk \setminus \operatorname{Spec}(gf)$, como queríamos. □

§6. El polinomio minimal

5.6.1. Proposición. *Sea V un espacio vectorial de dimensión finita y sea $f : V \to V$ un endomorfismo de V.*

(i) *Existe un único polinomio mónico $m_f \in \Bbbk[X]$ tal que*
  - $m_f(f) = 0$ *en* $\operatorname{End}(V)$ *y*
  - *cada vez que $p \in \Bbbk[X]$ es tal que $p(f) = 0$ en $\operatorname{End}(V)$, el polinomio m_f divide a p.*

(ii) *El subespacio $\langle f^i : i \in \mathbb{N}_0 \rangle$ de $\operatorname{End}(V)$ tiene dimensión igual al grado d de m_f y el conjunto $\{\mathrm{id}_V, f, \dots, f^{d-1}\}$ es una de sus bases. Si $a_0, \dots, a_{d-1} \in \Bbbk$ son los escalares tales que $f^d = a_0 \mathrm{id}_V + a_1 f + \cdots + a_{d-1} f^{d-1}$, entonces*

$$m_f = X^d - a_{d-1} X^{d-1} - \cdots - a_1 X - a_0.$$

El número d es el menor entero positivo r tal que $f^r \in \langle \mathrm{id}_V, f, \dots, f^{r-1}\rangle$.

Llamamos al polinomio m_f el ***polinomio minimal*** de f. La segunda parte de esta proposición nos da una cota para su grado —implica que éste no es mayor que $\dim \mathrm{End}(V) = (\dim V)^2$, ya que coincide con la dimensión de un subespacio de $\mathrm{End}(V)$— y, por otro lado, una forma de determinar explícitamente el polinomio m_f.

Demostración. Sea $\varepsilon_f : \Bbbk[X] \to \mathrm{End}(V)$ el homomorfismo de álgebras tal que $\varepsilon_f(X) = f$ y sea $I = \mathrm{Nu}(\varepsilon_f)$ su núcleo. Como $\mathrm{End}(V)$ tiene dimensión finita y $\Bbbk[X]$ no, la función ε_f no puede ser inyectiva. Esto implica que el ideal I de $\Bbbk[X]$ no es nulo y, de acuerdo a la Proposición 5.4.6, existe un único polinomio mónico $m_f \in I$ tal que para cada $p \in \Bbbk[X]$ se tiene que $p \in I$ si y solamente si p es divisible por m_f. Esto significa, en vista de la definición de I, que m_f tiene precisamente las dos propiedades descriptas en la parte (i) de la proposición.

Sea d el grado de m_f y sean $a_0, \dots, a_{d-1} \in \Bbbk$ tales que $m_f = X^d - a_{d-1}X^{d-1} - \cdots - a_1 X - a_0$. Como $m_f(f) = 0$, vale que

$$f^d = a_0 \mathrm{id}_V + a_1 f + \cdots + a_{d-1} f^{d-1}. \tag{14}$$

Consideremos los subespacios $S = \langle f^i : i \in \mathbb{N}_0\rangle$ y $F = \langle \mathrm{id}_V, f, \dots, f^{d-1}\rangle$ de $\mathrm{End}(V)$ y mostremos que, de hecho, $S = F$. Claramente $F \subseteq S$, así que basta probar que $S \subseteq F$ y, para ello, que $f^i \in F$ para todo $i \in \mathbb{N}_0$. Si $i = 0$ esto es evidente. Supongamos entonces que $i \geq 1$ e, inductivamente, que $f^{i-1} \in F$, de manera que existen escalares $b_0, \dots, b_{d-1} \in \Bbbk$ tales que $f^{i-1} = b_0 \mathrm{id}_V + b_1 d + \cdots + b_{d-1} f^{d-1}$. Componiendo a ambos lados de esta igualdad con f y usando (14), vemos entonces que

$$\begin{aligned} f^i &= b_0 f + b_1 f^2 + \cdots + b_{d-1} f^d \\ &= b_0 f + b_1 f^2 + \cdots + b_{d-1}(a_0 \mathrm{id}_V + a_1 f + \cdots + a_{d-1} f^{d-1}) \in F. \end{aligned}$$

Esto prueba lo que queremos.

Mostremos ahora que el conjunto $\mathscr{B} = \{\mathrm{id}_V, f, \dots, f^{d-1}\}$ es una base de F con d elementos y que, en particular, se tiene que $\dim S = \dim F = d$. Sean $b_0, \dots, b_{d-1} \in \Bbbk$ escalares tales que $b_0 \mathrm{id}_V + b_1 f + \cdots + b_{d-1} f^{d-1} = 0$. El polinomio $p = b_0 + b_1 X + \cdots + b_{d-1} X^{d-1} \in \Bbbk[X]$ es entonces tal que $p(f) = 0$ y es, de acuerdo a la parte (i) de la proposición, divisible por m_f. Como su grado es menor que el de m_f, esto implica que $p = 0$, esto es, que $b_0 = \cdots = b_{d-1} = 0$. Se sigue de esto que $\mathscr{B}$ tiene d elementos y que es linealmente independiente. En vista de cómo fue definido el subespacio F, vemos así que $\mathscr{B}$ es una base de F y que $\dim F = d$.

Finalmente, si r es un entero tal que $f^r \in \langle \mathrm{id}_V, f, \dots, f^{r-1}\rangle$, existen escalares $c_0, \dots, c_{r-1} \in \Bbbk$ tales que $f^r = a_0 \mathrm{id}_V + a_1 f + \cdots + a_{r-1} f^{r-1}$ y el polinomio $q = c_0 + c_1 X + \cdots + c_{r-1} X^{r-1} - X^r$ es tal que $q(f) = 0$. La parte (i) de la proposición nos dice entonces que m_f divide a q y, en consecuencia y porque q no es nulo, que $r \geq d$. □

5.6.2. Dos situaciones extremas de la Proposición 5.6.1 son las siguientes:

- Si V es un espacio nulo y $f : V \to V$ es un endomorfismo —necesariamente la función nula, claro— entonces el polinomio minimal de f es $m_f = 1$, que tiene grado nulo.

- Por otro lado, si V es un espacio de dimensión finita y positiva y $f : V \to V$ es el endomorfismo nulo se tiene que $m_f = X$.

5.6.3. Es importante observar que la hipótesis de que el espacio V tiene dimensión finita que aparece en la Proposición 5.6.1 es necesaria para obtener la conclusión. Por ejemplo, si

$$f : q \in \Bbbk[X] \mapsto Xq \in \Bbbk[X]$$

es el endomorfismo del espacio vectorial $\Bbbk[X]$ dado por la multiplicación por X, entonces para todo $p \in \Bbbk[X]$ no nulo se tiene que $p(f) \neq 0$: es fácil verificar que, de hecho, $p(f)(1) = p$. Esto implica que el homomorfismo de álgebras $\varepsilon_f : \Bbbk[X] \to \mathsf{End}(\Bbbk[X])$ es inyectivo en este caso y la demostración que hicimos no puede ni comenzar.

5.6.4. El polinomio minimal de un endomorfismo f de un espacio vectorial V de dimensión finita es, en el sentido preciso de la Proposición 5.6.1, el polinomio mónico p de menor grado tal que $p(f) = 0$. De una forma similar, podemos construir para cada vector x de V un polinomio caracterizado por ser el polinomio mónico p de menor grado tal que $p(f)$ se anula en x.

Proposición. *Sea V un espacio vectorial, sea $f : V \to V$ un endomorfismo de V y sea $x \in V$.*

(*i*) *El subespacio $\langle x\rangle_f$ de V generado por el conjunto $\{f^i(x) : i \in \mathbb{N}_0\}$ es f-invariante.*

En vista de esto, podemos considerar la restricción $f_x : \langle x\rangle_f \to \langle x\rangle_f$ de f a $\langle x\rangle_f$.

(*ii*) *Para cada $p \in \Bbbk[X]$ se tiene que $p(f_x) = 0$ si y solamente si $p(f_x)(x) = 0$.*

(*iii*) *Si $\langle x\rangle_f$ tiene dimensión finita, entonces existe un único polinomio mónico $m_{f,x} \in \Bbbk[X]$ tal que*
- *$m_{f,x}(f)(x) = 0$ en V, y*
- *si $p \in \Bbbk[X]$ es un polinomio tal que $p(f)(x) = 0$, entonces $m_{f,v}$ divide a p.*

El grado d de $m_{f,x}$ coincide con la dimensión de $\langle x\rangle_f$, el conjunto $\mathscr{B} = (x, f(x), \dots, f^{d-1}(x))$ es una base ordenada de $\langle x\rangle_f$ y la matriz de f_x con respecto a la base $\mathscr{B}$ es precisamente la matriz compañera de $m_{f,x}$,

$$[f_x]^{\mathscr{B}}_{\mathscr{B}} = C(m_{f,x}).$$

Llamamos al subespacio $\langle x\rangle_f$ de V el ***subespacio f-cíclico generado por x en V*** y al polinomio $m_{f,x}$ el ***polinomio minimal de f relativo a x***.

Demostración. Para probar la afirmación (*i*) es suficiente con mostrar que $f(f^i(x)) = f^{i+1}(x)$ está en $\langle x\rangle_f$ para cada $i \in \mathbb{N}_0$ y esto es evidente.

La necesidad de la condición dada en (*ii*) es clara, ya que por supuesto x pertenece al dominio de $p(f_x)$. Para ver la suficiencia, supongamos que $p \in \Bbbk[X]$ es tal que $p(f_x)(x) = 0$ y sea $y \in \langle x\rangle_f$, de manera que existen $n \in \mathbb{N}_0$ y escalares $b_0, \dots, b_n \in \Bbbk$ tales que $y = b_0x + b_1f(x) + \cdots + b_nf^n(x)$. El polinomio $q = b_0 + b_1X + \cdots + b_nX^n \in \Bbbk[X]$ es entonces tal que $q(f_x)(x) = y$ y, usando la Proposición 5.5.2, se tiene que

$$\begin{aligned} p(f_x)(y) &= p(f_x)\big(q(f_x)(x)\big) = \big(p(f_x) \circ q(f_x)\big)(x) = \big(q(f_x) \circ p(f_x)\big)(x) \\ &= q(f_x)\big(p(f_x)(x)\big) = 0. \end{aligned}$$

Vemos así que $p(f_x)$ se anula en todos los elementos de su dominio $\langle x\rangle_f$, esto es, que $p(f_x) = 0$. Esto prueba la parte (*ii*) de la proposición.

Supongamos ahora que el subespacio $\langle x\rangle_f$ tiene dimensión finita y sea $m_{f,x} \in \Bbbk[X]$ el polinomio minimal del endomorfismo f_x de $\langle x\rangle_f$. Como $m_{f,x}(f_x) = 0$, la Proposición 5.5.1 nos dice que

$$m_{f,x}(f)(x) = m_{f,x}(f_x)(x) = 0.$$

Por otro lado, si $p \in \Bbbk[X]$ es un polinomio tal que $p(f)(x) = 0$, tenemos, otra vez gracias a la Proposición 5.5.1, que $p(f_x)(x) = 0$ y esto, de acuerdo a la parte (*ii*) de la proposición, implica que $p(f_x) = 0$. De que $m_{f,x}$ sea el polinomio minimal de f_x, se sigue entonces que $m_{f,x}$ divide a p. Esto muestra que $m_{f,x}$ satisface las dos condiciones del enunciado. Si $\tilde{m} \in \Bbbk[X]$ es otro polinomio que las satisface, entonces $m_{f,x}$ y $\tilde{m}$ se dividen mutuamente y, como son mónicos, son necesariamente iguales.

Sea d el grado de $m_{f,x}$ y sean $a_0, \dots, a_{d-1} \in \Bbbk$ tales que

$$m_{f,x} = X^d + a_{d-1}X^{d-1} + \cdots + a_0.$$

Como $m_{f,x}(f)(x) = 0$, se tiene que

$$f^d(x) = -a_0x - a_1f(x) - \cdots - a_{d-1}f^{d-1}(x). \tag{15}$$

Consideremos el subespacio $F = \langle x, f(x), \dots, f^{d-1}(x)\rangle$ de V y mostremos que coincide con $\langle x\rangle_f$. La inclusión $F \subseteq \langle x\rangle_f$ es evidente y para ver la inclusión recíproca es suficiente que mostremos que $f^i(x) \in F$ para todo $i \in \mathbb{N}_0$. Cuando $i = 0$ esto es claro; por otro lado, si $i \geq 1$ y sabemos que $f^{i-1}(x) \in F$, esto es, que existen escalares $b_0, \dots, b_{d-1} \in \Bbbk$ tales que

$$f^{i-1}(x) = b_0x + b_1f(x) + \cdots + b_{d-1}f^{d-1}(x),$$

entonces

$$\begin{aligned} f^i(x) &= f(f^{i-1}(x)) = f(b_0x + b_1f(x) + \cdots + b_{d-1}f^{d-1}(x)) \\ &= b_0x + b_1f^2(x) + \cdots + b_{d-2}f^{d-1}(x) + b_{d-1}f^d(x) \in F, \end{aligned}$$

ya que los primeros $d-1$ sumandos de esta última suma están en F por la definición misma de F y el último por la igualdad (15).

Los vectores x, $f(x), \dots, f^{d-1}(x)$ de F son distintos dos a dos y linealmente independientes. En efecto, si no lo fueran habría escalares $c_0, \dots, c_{d-1} \in \Bbbk$ tales que $c_0x+c_1f(x)+\cdots+c_{d-1}f^{d-1}(x) = 0$ y, en consecuencia el polinomio $q = c_{d-1}X^{d-1} + \cdots + c_1X + c_0 \in \Bbbk[X]$ tendría la propiedad de que $q(f_x)(x) = q(f)(x) = 0$. Lo que sabemos de $m_{f,x}$, entonces, implicaría que $m_{f,x}$ lo divide y, como su grado es menor que el de $m_{f,x}$, esto sólo es posible si es nulo: así, debe ser $c_0 = \cdots = c_{d-1} = 0$.

Concluimos de esta forma que $\mathscr{B} = (x, f(x), \dots, f^{d-1}(x))$ es una base ordenada de F y, en particular, que la dimensión de F y, por consiguiente de $\langle x\rangle_f$, es exactamente d. Para completar la prueba de (*iii*) tenemos que determinar la matriz $[f_x]^{\mathscr{B}}_{\mathscr{B}}$. Como $f_x(f^i(x)) = f^{i+1}(x)$ para cada $i \in [\![0, d-1]\!]$ y

$$f_X(f^{d-1}(x)) = -a_0 x - a_1 f(x) - \cdots - a_{d-1} f^{d-1}(x),$$

se tiene que

$$[f_x]^{\mathscr{B}}_{\mathscr{B}} = \begin{pmatrix} 0 & 0 & 0 & \cdots & 0 & -a_0 \\ 1 & 0 & 0 & \cdots & 0 & -a_1 \\ 0 & 1 & 0 & \cdots & 0 & -a_2 \\ \vdots & \vdots & \ddots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \ddots & 0 & -a_{d-2} \\ 0 & 0 & 0 & \cdots & 1 & -a_{d-1} \end{pmatrix} = C(m_{f,x}),$$

la matriz compañera de m_{f_x}. □

5.6.5. Recordemos que decimos que un polinomio p de $\Bbbk[X]$ es el ***mínimo común múltiplo*** de una familia de polinomios $(p_i)_{i \in I}$ si es mónico,

- p es divisible por p_i para cada $i \in I$, y
- si un polinomio $q \in \Bbbk[X]$ es divisible por p_i para todo $i \in I$, entonces q es divisible por p.

Toda familia finita de elementos de $\Bbbk[X]$ tiene un mínimo común múltiplo, pero hay familias infinitas de polinomios que no admiten uno: por ejemplo, los polinomios de la familia $\{X^i : i \in \mathbb{N}_0\}$ no poseen ni siquiera un múltiplo común.

5.6.6. Proposición. *Sea V un espacio vectorial de dimensión finita y sea $f : V \to V$ un endomorfismo de V.*

(*i*) *El polinomio minimal m_f es el mínimo común múltiplo de los polinomios $m_{f,x}$ con $x \in V$.*

(*ii*) *Si $V_1, \dots, V_n$ son subespacios f-invariantes de V y $V = V_1 + \cdots + V_n$, entonces m_f es el mínimo común múltiplo de los polinomios minimales $m_{f_{V_1}}, \dots, m_{f_{V_n}}$ de las restricciones $f_{V_1}, \dots, f_{V_n}$.*

(*iii*) *Si $x_1, \dots, x_n$ son vectores de V tales que $V = \langle x_1 \rangle_f + \cdots + \langle x_n \rangle_f$, entonces m_f es el mínimo común múltiplo de los polinomios $m_{f,x_1}, \dots, m_{f,x_n}$.*

Demostración. (*i*) Si x es un vector de V, entonces $m_f(f_x)(x) = 0$, así que $m_{f,x}$ divide a m_f. Por otro lado, supongamos que $p \in \Bbbk[x]$ es un polinomio que es divisible por $m_{f,x}$ para cada $x \in V$. En ese caso, para cada $x \in V$ existe $q \in \Bbbk[X]$ tal que $p = q m_{f,x}$ y entonces

$$p(f)(x) = q(f)\big(m_{f,x}(f)(x)\big) = 0.$$

Vemos así que $p(f) = 0$ y la minimalidad de m_f implica que m_f divide a p.

(*ii*) Sean $V_1, \dots, V_n$ subespacios f-invariantes de V y supongamos que $V = V_1 + \cdots + V_n$. Si $i \in [\![n]\!]$, sabemos que $m_f(f_{V_i}) = m_f(f)_{V_i} = 0$, y entonces m_{V_i} divide a m_f: esto nos dice que m_f es un múltiplo común de los polinomios minimales $m_{f_{V_1}}, \dots, m_{f_{V_n}}$. Sea, por otro lado, $p \in \Bbbk[X]$ un múltiplo común de estos polinomios. Si $x \in V$, entonces existen $x_1 \in V_1, \dots, x_n \in V_n$ tales que $x = x_1 + \cdots + x_n$ y para cada $i \in [\![n]\!]$ es

$$p(f)(x_i) = p(f)_{V_i}(x_i) = p(f_{V_i})(x_i) = 0,$$

ya que $p(f_{V_i}) = 0$, y como consecuencia de esto tenemos que

$$p(f)(x) = p(f)(x_1) + \cdots + p(f)(x_n) = 0.$$

Esto nos dice que, de hecho, $p(f) = 0$ y, por lo tanto, que m_f divide a p.

(*iii*) Sean finalmente $x_1, \ldots, x_n$ vectores de V tales que $V = \langle x_1 \rangle_f + \cdots + \langle x_n \rangle_f$. Para cada $i \in [\![n]\!]$ sabemos de la Proposición 5.6.4 que el subespacio $\langle x_i \rangle_f$ es f-invariante y que el polinomio minimal de la restricción de f a $\langle x_i \rangle_f$ es $m_{f,x}$. El resultado de (*iii*) es consecuencia de esto y de la parte (*ii*) que ya probamos. □

5.6.7. La segunda parte de la Proposición 5.6.6 tiene la siguiente consecuencia útil:

Corolario. *Sea V un espacio vectorial de dimensión finita y sea $f : V \to V$ un endomorfismo de V. Si W es un subespacio f-invariante de V, entonces el polinomio minimal de la restricción $f_W : W \to W$ divide al de f.*

Demostración. En efecto, como V y W son subespacios f-invariantes de V y, por supuesto, $V = W + V$, la Proposición 5.6.6*(ii)* nos dice que el polinomio minimal m_f es el mínimo común múltiplo de los polinomios m_{f_W} y m_f y, en particular, es divisible por el primero de éstos. □

5.6.8. Por otro lado, un caso especial de la tercera parte de la Proposición 5.6.6 es muchas veces de interés:

Corolario. *Sea V un espacio vectorial de dimensión finita y sea $f : V \to V$ un endomorfismo de V. Si existe un vector $x \in V$ tal que $V = \langle x \rangle_f$, entonces $m_f = m_{f,x}$.*

Un vector x tal que $V = \langle x \rangle_f$, como en este corolario, es un ***vector cíclico para f***.

Demostración. Esto es el caso en el que $n = 1$ y $x_1 = x$ de la Proposición 5.6.6*(iii)*. □

5.6.9. El siguiente resultado es conocido como el *Teorema de Cayley–Hamilton*. Fue enunciado por Arthur Cayley en su *A memoir on the theory of matrices* [Cay58] de 1858 pero sólo para matrices de 2×2; Cayley menciona allí que sabe probarlo para matrices de 3×3 pero no lo hace y dice al pasar que «no le parece necesario emprender la prueba del teorema general». William Rowan Hamilton había probado antes, en su libro [Ham53] de 1853 sobre los cuaterniones, que una matriz de 3×3 satiface a un polinomio de grado 3 y que una de 4×4 satisface uno de grado 4, pero no identificó ese polinomio con el polinomio característico. La primera prueba completa del teorema general fue dada por Frobenius en 1978.

Teorema. *Sea V un espacio vectorial de dimensión finita y sea $f : V \to V$ un endomorfismo de V. Si χ_f es el polinomio característico de f, entonces $\chi_f(f) = 0$ y, en particular, χ_f es divisible por el polinomio minimal m_f.*

Demostración. Sea $x \in V$ y sea W un complemento del subespacio cíclico $\langle x \rangle_f$ en V, de manera que $V = \langle x \rangle_f \oplus W$. Si $d = \dim \langle x \rangle_f$, sabemos que $\mathscr{B}_x = (x, f(x), \ldots, f^{d-1}(x))$ es una base ordenada de $\langle x \rangle_f$; sea, por otro lado, $m = \dim W$ y sea $(y_1, \ldots, y_m)$ una base ordenada de W. En esta situación sabemos que $\mathscr{B} = (x, f(x), \ldots, f^{d-1}(x), y_1, \ldots, y_m)$ es una base ordenada de V y, como $\langle x\rangle_f$ es un subespacio f-invariante, es inmediato ver que la matriz de f con respecto a esta base ordenada $\mathscr{B}$ tiene una descomposición en bloques de la forma

$$[f]_{\mathscr{B}}^{\mathscr{B}} = \begin{pmatrix} [f_x]_{\mathscr{B}_x}^{\mathscr{B}_x} & C \\ 0 & D \end{pmatrix}$$

para ciertas matrices $D \in \mathsf{M}_m(\Bbbk)$ y $C \in \mathsf{M}_{d,m}(\Bbbk)$. El polinomio característico de f es, entonces,

$$\chi_f = \chi_{f_x} \cdot \chi_C = m_{f,x} \cdot \chi_C.$$

Vemos así que el polinomio característico χ_f es divisible por cada uno de los polinomio $m_{f,x}$ con $x \in V$ y, en vista de la Proposición 5.6.6*(iii)*, que es divisible por m_f. Esto implica, como sabemos, que $\chi_f(f) = 0$, como afirma el teorema. □

5.6.10. Es una consecuencia de la Proposición 5.6.1 que el grado del polinomio minimal de un endomorfismo $f : V \to V$ de un espacio vectorial de dimensión n es a lo sumo n^2, ya que coincide con la dimensión de un cierto subespacio de $\mathsf{End}(V)$. Del teorema de Cayley-Hamilton obtenemos trivialmente una cota mucho mejor:

Corolario. *Sea V un espacio vectorial de dimensión finita n y sea $f : V \to V$ un endomorfismo de V. El polinomio minimal de f tiene grado a lo sumo igual a n. Si ese grado es igual a n, entonces el polinomio minimal de f coincide con el polinomio característico de f.*

Demostración. La primera afirmación es consecuencia inmediata de que el polinomio minimal divide al característico y de que este último tiene grado n. La segunda, de que si los dos polinomios tienen el mismo grado que uno divida al otro implica —como son mónicos— que son iguales. □

5.6.11. Del teorema se deduce fácilmente el siguiente resultado, que es muchas veces útil para determinar los autovalores de un endomorfismo:

Corolario. *Sea V un espacio vectorial de dimensión finita y sea $f : V \to V$ un endomorfismo. Los polinomios minimal y característico de f tienen las mismas raíces en $\Bbbk$ y, en particular, todo autovalor de f es raíz del polinomio minimal de f.*

Más adelante, en la Proposición 5.7.4, daremos una mejora de este corolario.

Demostración. Si $\lambda \in \Bbbk$ es una raíz del polinomio característico χ_f, sabemos que existe un autovector $x \in V$ de autovalor λ. Esto implica, claramente, que $m_{f,x} = X - \lambda$, y la Proposición 5.6.6*(i)* nos dice que este polinomio divide al polinomio minimal m_f, así que λ es una raíz de m_f.

Recíprocamente, como el Teorema de Cayley–Hamilton nos dice que m_f divide a χ_f, toda raíz del primero es raíz del segundo. □

5.6.12. Una tercera consecuencia directa del teorema de Cayley–Hamilton es la siguiente expresión para el endomorfismo inverso de un endomorfismo inversible:

Corolario. *Sea V un espacio vectorial de dimensión finita n, sea $f : V \to V$ un endomorfismo y supongamos que el polinomio característico de f es $\chi_f = X^n + a_1X^{n-1} + \cdots + a_{n-1}X + a_n$. Si f es* *inversible, entonces* $a_n \neq 0$ *y*

$$f^{-1} = -a_n^{-1} f^{n-1} - a_n^{-1} a_1 f^{n-2} - \cdots - a_n^{-1} a_{n-1} \mathsf{id}_V.$$

Demostración. Del teorema de Cayley–Hamilton sabemos que $\chi_f(f) = 0$ y entonces

$$f^n + a_1 f^{n-1} + \cdots + a_{n-1} f + a_n \mathsf{id}_V = 0.$$

Si f es inversible, la Proposición 4.10.2*(iii)* nos dice que $\det f \neq 0$ y, de acuerdo a la Proposición 5.3.7, es entonces $a_n \neq 0$. Componiendo a la derecha, por ejemplo, con f^{-1} a ambos lados de esta igualdad y multiplicando por a_n^{-1}, entonces, vemos que

$$f^{-1} = -a_n^{-1} f^{n-1} - a_n^{-1} a_1 f^{n-2} - \cdots - a_n^{-1} a_{n-1} \mathsf{id}_V,$$

como afirma la proposición. □

§7. Descomposición primaria y diagonalizabilidad

5.7.1. El siguiente resultado nos provee de la llamada *descomposición primaria* de un espacio vectorial con respecto a uno de sus endomorfismos:

Proposición. *Sea V un espacio vectorial de dimensión finita, sea $f : V \to V$ un endomorfismo de V y sea m_f su polinomio minimal. Si $p_1, \dots, p_n \in \Bbbk[X]$ son polinomios mónicos no constantes coprimos dos a dos tales que $m_f = p_1 \cdots p_n$ y ponemos $V_i = \mathsf{Nu}(p_i(f))$ para cada $i \in [\![n]\!]$, entonces tenemos una descomposición en suma directa*

$$V = V_1 \oplus \cdots \oplus V_n. \tag{16}$$

Para cada $i \in [\![n]\!]$ el subespacio V_i es no nulo y f-invariante, el polinomio minimal de la restricción $f_{V_i} : V_i \to V_i$ de f a V_i es precisamente p_i, y si $\pi_i : V \to V$ es el proyector correspondiente a V_i para la descomposición en suma directa (16) de V, entonces hay un polinomio $v_i \in \Bbbk[X]$ tal que $\pi_i = v_i(f)$ y, en particular, π_i conmuta con f.

Llamamos a la descomposición de V que nos da esta proposición la ***descomposición primaria*** de V relativa a f y el subespacio $V_i = \mathsf{Nu}(p_i(f))$ es la ***componente p-primaria*** de V.

Demostración. Sean $p_1, \dots, p_n \in \Bbbk[X]$ polinomios no constantes coprimos dos a dos tales que $m_f = p_1 \cdots p_n$, como en el enunciado, y para cada $i \in [\![n]\!]$ consideremos el polinomio

$$q_i = p_1 \cdots \hat{p}_i \cdots p_n,$$

producto de todos ellos salvo el i-ésimo. Como los polinomios $p_1, \dots, p_n$ son coprimos dos a dos,

los polinomios $q_1, \dots, q_n$ no poseen divisores no triviales comunes. (17)

Para probar esto, supongamos que $t \in \Bbbk[X]$ es un divisor común no trivial de esos polinomios y, sin pérdida de generalidad, que es irreducible. Como t divide a q_1, existe $k \in [\![2, r]\!]$ tal que t divide a p_k, y como t también divide a q_k, existe $s \in [\![n]\!] \setminus \{k\}$ tal que t divide a p_s: esto es absurdo, ya que p_k y p_s no tienen divisores comunes.

Como consecuencia de (17) existen polinomios $r_1, \dots, r_n \in \Bbbk[X]$ tales que

$$r_1 q_1 + \cdots + r_n q_n = 1. \tag{18}$$

Para cada $i \in [\![n]\!]$ sea $\pi_i = r_i(f) q_i(f) : V \to V$. Afirmamos que

$$\pi_1 + \cdots + \pi_n = \mathsf{id}_V \tag{19}$$

y que para cada $i, j \in [\![n]\!]$ es

$$\pi_i \circ \pi_j = \begin{cases} \pi_i, & \text{si } i = j; \\ 0, & \text{en caso contrario.} \end{cases} \tag{20}$$

Probemos esto:

- La igualdad (19) es consecuencia inmediata de (18) y de la definición de los endomorfismos $\pi_1, \dots, \pi_n$.
- Si i y j son dos elementos distintos de $[\![n]\!]$, entonces p_j divide a q_i, así que existe $s \in \Bbbk[X]$ con $q_i = s p_j$, y por lo tanto el polinomio $r_i q_i r_j q_j$ es divisible por m_f, ya que es igual a $r_i r_j s p_j q_j$. Como consecuencia de esto,

  $$\pi_i \circ \pi_j = (r_i q_i r_j q_j)(f) = 0.$$

- Por otro lado, si $i \in [\![n]\!]$, componiendo ambos miembros de la igualdad (19) con π_i vemos que

  $$\pi_i \circ \pi_1 + \cdots + \pi_i \circ \pi_n = \pi_i$$

  y, en vista de lo que ya probamos, el miembro izquierdo de esta ecuación es igual a $\pi_i \circ \pi_i$. Esto nos dice que $\pi_i \circ \pi_i = \pi_i$ y prueba (20).

Ahora bien, como consecuencia de (19) y (20), la Proposición 2.9.1 nos dice que hay una descomposición en suma directa

$$V = \mathsf{Im}(\pi_1) \oplus \cdots \oplus \mathsf{Im}(\pi_n)$$

cuyos proyectores asociados son precisamente los endomorfismos $\pi_1, \dots, \pi_n$, cada uno de los cuales se obtiene evaluando un polinomio de $\Bbbk[X]$ en f.

Mostremos ahora que para cada $i \in [\![n]\!]$ se tiene que

$$\mathsf{Nu}(p_i(f)) = \mathsf{Im}(\pi_i). \tag{21}$$

- Si x es un vector de en $\mathsf{Im}(\pi_i)$, existe $y \in V$ tal que $x = \pi_i(y) = (r_i q_i)(f)(y)$ y entonces

$$p_i(f)(x) = (p_i r_i q_i)(f)(y) = (r_i m_f)(f)(y) = 0,$$

  es decir, x pertenece a $\mathsf{Nu}(p_i i(f))$
- Recíprocamente, si $x \in V_i = \mathsf{Nu}(p_i(f))$, se tiene que para cada para cada $j \in [\![n]\!] \setminus \{i\}$ es $q_j(f)(x) = 0$ y, en consecuencia, $\pi_j(x) = 0$: recordando la igualdad (19), tenemos entonces que $x = \pi_1(x) + \cdots + \pi_n(x) = \pi_i(x) \in \mathsf{Im}(\pi_i)$. Esto prueba (21).

Para cada $i \in [\![n]\!]$ escribamos V_i al subespacio $\mathsf{Nu}(p_i(f))$, de manera que

$$V = V_1 \oplus \cdots \oplus V_n. \tag{22}$$

Como el endomorfismo $p_i(f)$ conmuta con f, para cada $x \in V_i$ es

$$p_i(f)\big(f(x)\big) = f\big(p_i(f)(x)\big) = f(0) = 0,$$

de manera que $f(x)$ también está en V_i: esto nos dice que V_i es f-invariante. Podemos entonces considerar la restricción $f_{V_i} : V_i \to V_i$. Como

$$p_i(f_{V_i}) = p_i(f)_{V_i} = 0,$$

el polinomio $m_{f_{V_i}}$ divide a p_i y, por lo tanto, existe un polinomio mónico $s_i \in \Bbbk[X]$ tal que

$$p_i = s_i m_{f_{V_i}}. \tag{23}$$

Se sigue de esto, en particular, que

los polinomios $m_{f_{V_1}}, \dots, n_{f_{V_n}}$ *son coprimos dos a dos* ,

ya que los polinomios $p_1, \dots, p_n$ lo son por hipótesis. Usando esto y la Proposición 5.6.6*(ii)* aplicada a la descomposición (22), vemos que

$$m_{f_{V_1}} \cdots m_{f_{V_n}} = \mathsf{mcm}\{m_{f_{V_1}}, \dots, m_{f_{V_n}}\} = m_f. \tag{24}$$

Por otro lado, de (23) se sigue que

$$m_f = p_1 \cdots p_n = s_1 \cdots s_n m_{f_{V_1}} \cdots m_{f_{V_n}}$$

y de comparar esto con (24) es claro que el polinomio $s_1 \cdots s_n$ debe ser igual a 1. Esto implica que $m_{f_{V_i}} = p_i$ para cada $i \in [\![n]\!]$ y, en particular, que V_i es un subespacio no nulo de V. Con esto todas las afirmaciones de la proposición quedan probadas. □

5.7.2. Usando este resultado sobre descomposiciones primarias, podemos dar una condición necesaria y suficiente para la diagonalizabilidad de un endomorfismo en términos de su polinomio minimal, que es el resultado central de este capítulo:

Teorema. *Sea V un espacio vectorial de dimensión finita positiva y sea $f : V \to V$ un endomorfismo de V. El endomorfismo f es diagonalizable si y solamente si su polinomio minimal m_f se factoriza como producto de polinomios lineales en $\Bbbk[X]$ y todas sus raíces son simples.*

Demostración. Supongamos primero que f es diagonalizable y sean $\lambda_1, \ldots, \lambda_n$ los autovalores de f listados sin repeticiones. De acuerdo a la Proposición 5.2.1, tenemos una descomposición

$$V = E_{\lambda_1}(f) \oplus \cdots \oplus E_{\lambda_n}(f)$$

de V como suma directa de subespacios no nulos y f-invariantes. Es claro que el polinomio minimal de la restricción $f_{E_{\lambda_i}(f)}$ de f al autoespacio $E_{\lambda_i}(f)$ es el polinomio $X - \lambda_i$ y entonces la Proposición 5.6.6*(ii)* nos dice que el polinomio minimal de f es el mínimo común múltiplo de los n polinomios $X - \lambda_1, \ldots, X - \lambda_n$, que es claramente

$$(X - \lambda_1) \cdots (X - \lambda_n).$$

Este polinomio manifiestamente se factoriza en $\Bbbk[X]$ como producto de polinomio lineales y tiene sus raíces simples: esto significa que la condición del enunciado es necesaria.

Veamos su suficiencia. Supongamos que el polinomio minimal de f posee en $\Bbbk[X]$ una factorización de la forma

$$m_f = (X - \lambda_1) \cdots (X - \lambda_n)$$

con $\lambda_1, \ldots, \lambda_n \in \Bbbk$ distintos dos a dos. Como los polinomios $X - \lambda_1, \ldots, X - \lambda_n$ son mónicos no constantes y no tiene divisores comunes, la Proposición 5.7.1 nos dice que si para cada $i \in [\![n]\!]$ ponemos $V_i = \mathsf{Nu}(f - \lambda \mathsf{id}_V)$, entonces

$$V = V_1 \oplus \cdots \oplus V_n.$$

Para cada $i \in [\![n]\!]$ se tiene que $V_i = E_{\lambda_i}(f)$, el autoespacio de f correspondiente a λ_i, así que la existencia de esta descomposición, en vista de la Proposición 5.2.1, implica que el endomorfismo f es diagonalizable. □

5.7.3. Ejemplo. Sea V un espacio vectorial complejo de dimensión finita y sea $f : V \to V$ un endomorfismo de V. Decimos que f tiene ***orden finito*** si existe un entero positivo $k \in \mathbb{N}$ tal que $f^k = \mathsf{id}_V$ y en ese caso llamamos al entero k el ***orden*** de f. Afirmamos que

si f tiene orden finito, entonces f es diagonalizable.

En efecto, si f tiene orden finito k, de manera que $f^k - \mathsf{id}_V = 0$, y $p = X^k - 1 \in \Bbbk[X]$, entonces $p(f) = 0$. Esto significa que el polinomio minimal m_f divide a p y, como éste último tiene raíces simples, aquél también. Como m_f se factoriza como producto de factores lineales porque el cuerpo $\mathbb{C}$ es algebraicamente cerrado, el Teorema 5.7.2 nos dice que f es diagonalizable. ◇

5.7.4. Una segunda consecuencia de la existencia de descomposiciones primarias es:

Proposición. *Sea V un espacio vectorial de dimensión finita y sea $f : V \to V$ un endomorfismo de V. Los polinomios minimal y característico de f tienen los mismos divisores irreducibles.*

Observemos que esto completa el resultado del Corolario 5.6.11, que afirma lo mismo pero sólo sobre los divisores de grado 1.

Demostración. El Teorema de Cayley–Hamilton 5.6.9 nos dice que m_f divide a χ_f, así que si un polinomio irreducible divide al primero también divide al segundo.

Supongamos ahora que $p_1, \dots, p_n$ son polinomios mónicos irreducibles distintos dos a dos y que $\nu_1, \dots, \nu_n$ son enteros positivos tales que $m_f = p_1^{\nu_1} \cdots p_n^{\nu_n}$. Como los polinomios $p_1^{\nu_1}, \dots, p_n^{\nu_n}$ son mónicos, no constantes y coprimos dos a dos, la Proposición 5.7.1 nos dice que si para cada $i \in [\![n]\!]$ ponemos $V_i = \mathsf{Nu}(p_i^{\nu_i}(f))$, entonces hay una descomposición $V = V_1 \oplus \cdots \oplus V_n$ de V como suma directa de subespacios no nulos y f-invariantes. Para cada $i \in [\![n]\!]$ podemos considerar la restricciones $f_{V_i} : V_i \to V_i$ de f a V_i y el Ejemplo 5.3.6(a) nos dice que $\chi_f = \chi_{f_{V_1}} \cdots \chi_{f_{V_n}}$. Si $i \in [\![n]\!]$, sabemos de la Proposición 5.7.1 que el polinomio minimal de la restricción f_{V_i} es $m_{f_{V_i}} = p_i^{\nu_i}$ y del Teorema de Cayley–Hamilton que $m_{f_{V_i}}$ divide a $\chi_{f_{V_i}}$: todo esto implica, por supuesto, que p_i divide a χ_f. □

§8. Endomorfismos triangularizables y semisimples

5.8.1. El Teorema 5.7.2 nos dice que un endomorfismo $f : V \to V$ de un espacio vectorial de dimensión finita es diagonalizable si en la factorización de su polinomio minimal como producto de factores irreducibles mónicos

- todos los factores tienen grado 1, y
- en esa factorización no aparecen factores repetidos.

En esta sección nos proponemos analizar qué sucede cuando el polinomio minimal sólo satisface a una de estas condiciones.

5.8.2. Empecemos por el más simple de los dos resultados:

Proposición. *Sea V un espacio vectorial de dimensión finita y sea $f : V \to V$ un endomorfismo de V. El polinomio minimal de f se factoriza como producto de factores lineales si y solamente si existe una base ordenada $\mathscr{B}$ de V tal que la matriz $[f]^{\mathscr{B}}_{\mathscr{B}}$ es una matriz triangular superior.*

Decimos en ese caso, por razones obvias, que el endomorfismo f es ***triangularizable***. Observemos que la Proposición 5.3.13, que tiene una conclusión similar, se diferencia de ésta en que allí pusimos una hipótesis fuerte sobre el cuerpo de base mientras que aquí la hipótesis importante es sobre el endomorfismo mismo.

Demostración. HACER. □

5.8.3. Para poder describir qué significa que el polinomio minimal de un endomorfismo no tenga factores repetidos en su factorización como producto de polinomio irreducibles necesitamos algunos preliminares.

Proposición.

(i) *Sea A un álgebra, sea I un ideal de A y sea B el espacio cociente A/I. Existe un función bilineal $\cdot : B \times B \to \mathscr{B}$ tal que*

$$[a] \cdot [a'] = [aa']$$

cada vez que $a, a' \in A$ y con esta multiplicación B resulta ser un álgebra. Su elemento identidad es $[1_A]$ y la proyección canónica $\pi : A \to B$ es un homomorfismo de álgebras.

(ii) *Sea p un elemento irreducible de $\Bbbk[X]$ y sea (p) el ideal que genera en $\Bbbk[X]$. El álgebra $\Bbbk[X]/(p)$ es un cuerpo.*

Demostración. **Hacer.** □

5.8.4. Decimos que un endomorfismo $f : V \to V$ de un espacio vectorial V es ***semisimple*** si todo subespacio f-invariante de V posee un complemento f-invariante, esto es, si cada vez que W es un subespacio f-invariante de V existe otro subespacio f-invariante W' tal que $V = W \oplus W'$.

Proposición. *Sea V un espacio vectorial de dimensión finita y sea $f : V \to V$ un endomorfismo de V. El endomorfismo f es semisimple si y solamente si su polinomio minimal m_f es libre de cuadrados.*

Demostración. **Hacer.** □

Capítulo 6

Formas normales

§1. Equivalencia de funciones lineales y de matrices

6.1.1. Sean V y W dos espacios vectoriales. Decimos que dos funciones lineales $f, g : V \to W$ son ***equivalentes*** si hay automorfismos $\alpha : V \to V$ y $\beta : W \to W$ tales que $f = \beta \circ g \circ \alpha$, y en ese caso escribimos $f \sim g$. Esta relación en el conjunto $\hom(V, W)$ de todos las funciones lineales $V \to W$ es una relación de equivalencia:

- Si $f \in \hom(V, W)$, entonces por supuesto es $f = \mathrm{id}_W \circ f \circ \mathrm{id}_V$ y, como las funciones id_V e id_W son automorfismos de V y de W, respectivamente, se tiene que $f \sim f$.
- Si $f, g \in \hom(V, W)$ son tales que $f \sim g$, entonces existen automorfismos $\alpha : V \to V$ y $\beta : W \to W$ tales que $f = \beta \circ g \circ \alpha$ y, por lo tanto, $g = \beta^{-1} \circ f \circ \alpha^{-1}$, de manera que $g \sim f$.
- Si $f, g, h \in \hom(V, W)$ son tales que $f \sim g$ y $g \sim h$ y $\alpha, \alpha' : V \to V$ y $\beta, \beta' : W \to W$ son automorfismos tales que $f = \beta \circ g \circ \alpha$ y $g = \beta' \circ h \circ \alpha'$, entonces $f = \beta\beta' \circ h \circ \alpha'\alpha$. Como las composiciones $\alpha'\alpha$ y $\beta\beta'$ son automorfismos de V y de W, vemos así que $f \sim h$.

6.1.2. La relación de equivalencia de funciones lineales tiene una descripción bien concreta en término de matrices:

Proposición. *Sean V y W dos espacios vectoriales de dimensión finita. Dos funciones lineales $f, g : V \to W$ son equivalentes si y solamente si existen bases $\mathscr{B}_1$, $\mathscr{B}_1'$ de V y $\mathscr{B}_2$, $\mathscr{B}_2'$ de W tales que*

$$[f]^{\mathscr{B}_1}_{\mathscr{B}_2} = [g]^{\mathscr{B}_1'}_{\mathscr{B}_2'}. \tag{1}$$

Demostración. Sean $f, g : V \to W$ funciones lineales y sean n y m las dimensiones de V y de W, respectivamente. Para ver la suficiencia de la condición, supongamos que f y g son equivalentes y sean $\alpha : V \to V$ y $\beta : W \to W$ automorfismos tales que $f = \beta \circ g \circ \alpha$. Fijemos además bases ordenadas $\mathscr{B}_1 = (x_1, \dots, x_n)$ y $\mathscr{B}_2 = (y_1, \dots, y_m)$ de V y de W. Como α y β^{-1} son automorfismos, $\mathscr{B}_1' = (\alpha(x_1), \dots, \alpha(x_n))$ y $\mathscr{B}_2' = (\beta^{-1}(y_1), \dots, \beta^{-1}(y_m))$ también son bases ordenadas de V y de W, y podemos considerar la matriz $[g]^{\mathscr{B}_1'}_{\mathscr{B}_2'} = (a_{i,j}) \in \mathsf{M}_{m,n}(\Bbbk)$. Esto significa que para cada

$i \in [\![n]\!]$ es

$$g(\alpha(x_i)) = a_{1,i}\beta^{-1}(y_1) + \cdots + a_{m,1}\beta^{-1}(y_m)$$

y entonces, aplicando la función β a ambos lados de esta igualdad, que

$$f(x_i) = \beta(g(\alpha(x_i))) = a_{1,i}y_1 + \cdots + a_{m,1}y_m.$$

Vemos así que $[f]^{\mathscr{B}_1}_{\mathscr{B}_2} = (a_{i,j})$ y, en definitiva, que vale la igualdad (1) del enunciado.

Mostremos ahora la suficiencia de la condición. Supongamos que hay bases ordenadas $\mathscr{B}_1 = (x_1, \dots, x_n)$ y $\mathscr{B}'_1 = (x'_1, \dots, x'_n)$ de V y $\mathscr{B}_2 = (y_1, \dots, y_m)$ y $\mathscr{B}'_2 = (y'_1, \dots, y'_m)$ de W tales que la igualdad (1) del enunciado se cumple. Hay funciones lineales $\alpha : V \to V$ y $\beta : W \to W$ tales que $\alpha(x_i) = x'_i$ para cada $i \in [\![n]\!]$ y $\beta(y'_j) = y_j$ para cada $j \in [\![m]\!]$, y estas funciones son isomorfismos. Como claramente

Hacer: probar que si una función manda una base a una base es un iso!

$$[\alpha]^{\mathscr{B}_1}_{\mathscr{B}'_1} = I_n, \qquad [\beta]^{\mathscr{B}'_2}_{\mathscr{B}_2} = I_m,$$

usando la Proposición 2.8.5 y la hipótesis vemos que

$$[\beta \circ g \circ \alpha]^{\mathscr{B}_1}_{\mathscr{B}_2} = [\beta]^{\mathscr{B}'_2}_{\mathscr{B}_2} \cdot [g]^{\mathscr{B}'_1}_{\mathscr{B}'_2} \cdot [\alpha]^{\mathscr{B}_1}_{\mathscr{B}'_1} = I_m \cdot [g]^{\mathscr{B}'_1}_{\mathscr{B}'_2} \cdot I_n = [g]^{\mathscr{B}'_1}_{\mathscr{B}'_2} = [f]^{\mathscr{B}_1}_{\mathscr{B}_2}.$$

Así, las funciones lineales $\beta \circ g \circ \alpha$ y f tienen las mismas matrices con respecto a las bases $\mathscr{B}_1$ y $\mathscr{B}_2$, y esto implica que son ellas mismas iguales, de manera que las funciones f y g son equivalentes. □

6.1.3. La proposición que sigue resuelve de una manera satisfactoria el problema de decidir cuándo dos funciones lineales son equivalentes, ya que lo reduce al cálculo de los rangos de esas funciones.

Proposición. *Sean V y W espacios vectoriales de dimensión finita n y m, respectivamente.*

(*i*) *Si $f : V \to W$ es una función lineal de rango r, entonces existen bases $\mathscr{B}$ de V y $\mathscr{B}'$ de W tales que la matriz $[f]^{\mathscr{B}}_{\mathscr{B}'}$ es la matriz de bloques*

$$\begin{array}{cc} & \begin{array}{cc} r & n-r \end{array} \\ \begin{array}{c} r \\ m-r \end{array} & \begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix} \end{array}$$

(*ii*) *Dos funciones lineales $V \to W$ son equivalentes si y solamente si tienen el mismo rango*

(*iii*) *El número de clases de equivalencia de elementos de* $\hom(V, W)$ *es* $\min\{\dim V, \dim W\} + 1$.

Demostración. (*i*) Sea $f : V \to W$ una función lineal de rango r, de manera que $\dim \mathrm{Im}(f) = r$. Del teorema de la dimensión 2.5.1 sabemos que

$$\dim \mathrm{Nu}(f) = \dim V - \dim \mathrm{Im}(f) = n - r,$$

y entonces existe una base ordenada $\mathscr{B} = (x_1, \dots, x_n)$ de V tal que $(x_{r+1}, \dots, x_n)$ es una base de $\mathrm{Nu}(f)$. Mostremos que que $(f(x_1), \dots, f(x_r))$ es una base del subespacio $\mathrm{Im}(f)$ de W:

- Si $y \in \mathsf{Im}(f)$, entonces existe $x \in V$ tal que $y = f(x)$ y, como $\mathscr{B}$ es una base de V, hay escalares $a_1, \dots, a_n \in \Bbbk$ tales que $x = a_1x_1 + \cdots + a_nx_n$. Aplicando la función f a ambos lados de esta igualdad y recordando que los vectores $x_{r+1}, \dots, x_n$ están en el núcleo de f, vemos que $y = f(x) = a_1f(x_1) + \cdots + a_rf(x_r))$. Esto muestra que $\mathsf{Im}(f) = \langle f(x_1), \dots, f(x_r)\rangle$.
- Por otro lado, supongamos que $b_1, \dots, b_r \in \Bbbk$ son escalares tales que

  $$b_1f(x_1) + \cdots + b_rf(x_r) = 0.$$

  Esto implica que $f(b_1x_1 + \cdots + b_rx_r) = 0$ y, por lo tanto, que el vector $b_1x_1 + \cdots + b_rx_r$ está en $\mathsf{Nu}(f)$. En vista de la forma en que elegimos la base $\mathscr{B}$, entonces, existen escalares $c_{r+1}, \dots, c_n \in \Bbbk$ tales que

  $$b_1x_1 + \cdots + b_rx_r = c_{r+1}x_{r+1} + \cdots + c_rx_n$$

  y, como el conjunto $\mathscr{B}$ es linealmente independiente, esto implica que $b_1 = \cdots = b_r = 0$.

Ahora, como $(f(x_1), \dots, f(x_r))$ es una base de $\mathsf{Im}(f)$, podemos elegir vectores $y_1, \dots, y_{n-r} \in W$ de manera tal que $\mathscr{B}' = (f(x_1), \dots, f(x_r), y_1, \dots, y_{n-r})$ sea una base ordenada de W. Gracias a todas estas elecciones, es inmediato verificar que la matriz de f con respecto a las bases ordenadas $\mathscr{B}$ y $\mathscr{B}'$ es la matriz de bloques que aparece en el enunciado.

(*ii*) Sean $f, g : V \to W$ dos funciones lineales. Si f y g tienen el mismo rango r, entonces la primera parte de la proposición nos dice que existen bases ordenadas $\mathscr{B}_1$ y $\mathscr{B}_1'$ de V y $\mathscr{B}_2$ y $\mathscr{B}_2'$ de W tales que las matrices $[f]^{\mathscr{B}_1}_{\mathscr{B}_2}$ y $[g]^{\mathscr{B}_1'}_{\mathscr{B}_2'}$ son iguales a la matriz de bloques que aparece en su enunciado. Se sigue, claro, que estas dos matrices son iguales entre sí y, de acuerdo a la Proposición 6.1.2, que las funciones f y g son equivalentes.

Recíprocamente, si f y g son equivalentes, esa proposición nos dice que existen bases ordenadas $\mathscr{B}_1$ y $\mathscr{B}_1'$ de V y $\mathscr{B}_2$ y $\mathscr{B}_2'$ de W tales que $[f]^{\mathscr{B}_1}_{\mathscr{B}_2} = [g]^{\mathscr{B}_1'}_{\mathscr{B}_2'}$ y, como el rango de f coincide con el de la matriz $[f]^{\mathscr{B}_1}_{\mathscr{B}_2}$ y el de f con el de $[g]^{\mathscr{B}_1'}_{\mathscr{B}_2'}$, es claro que f y g tienen el mismo rango.

(*iii*) Sea $s = \min\{\dim V, \dim W\}$. Si $f : V \to W$ es una función lineal, entonces

$$0 \leq \dim \mathsf{Im}(f) \leq s,$$

de manera que hay $s+1$ posibles valores para $\dim \mathsf{Im}(f)$. Como dos funciones lineales con el mismo rango son equivalentes, esto nos dice que hay a lo sumo $s+1$ clases de equivalencia en $\mathsf{hom}(V, W)$. Para ver que son exactamente $s+1$ es suficiente con mostrar que para todo $r \in [\![0, s]\!]$ existe alguna función $V \to W$ de rango r. Esto es fácil: si $\mathscr{B} = (x_1, \dots, x_n)$ y $\mathscr{B}' = (y_1, \dots, y_m)$, entonces sabemos que hay una función lineal $f : V \to W$ tal que

$$f(x_i) = \begin{cases} y_i, & \text{si } i \leq r; \\ 0, & \text{en caso contrario,} \end{cases}$$

y es inmediato que $\dim \mathsf{Im}(f) = r$. □

6.1.4. Como con todo, la noción de equivalencia de funciones lineales se transpone al contexto de las matrices. Si m, $n \in \mathbb{N}$, decimos que dos matrices A, $B \in \mathsf{M}_{m,n}(\Bbbk)$ son ***equivalentes*** si existen matrices inversibles $P \in \mathsf{GL}_m(\Bbbk)$ y $Q \in \mathsf{GL}_n(\Bbbk)$ tales que $A = PBQ$ y en ese caso escribimos $A \sim B$. Un argumento similar al de 6.1.1 muestra que esto define una relación de equivalencia en el conjunto $\mathsf{M}_{m,n}(\Bbbk)$.

Lema. *Sean m, $n \in \mathbb{N}$ y sean A, $B \in \mathsf{M}_{m,n}(\Bbbk)$. Las siguientes afirmaciones son equivalentes:*

(a) *Las matrices A y B son equivalentes.*

(b) *Las funciones lineales $f_A : x \in \Bbbk^n \mapsto Ax \in \Bbbk^m$ y $f_B : x \in \Bbbk^n \mapsto Bx \in \Bbbk^m$ son equivalentes.*

Demostración. Si A y B son equivalentes, entonces existen matrices inversibles $P \in \mathsf{GL}_m(\Bbbk)$ y $Q \in \mathsf{GL}_n(\Bbbk)$ tales que $A = PBQ$. Las funciones $\alpha : x \in \Bbbk^n \mapsto Qx \in \Bbbk^n$ y $\beta : x \in \Bbbk^m \mapsto Px \in \Bbbk^m$ son entonces automorfismos y es inmediato verificar que $f_A = \beta \circ f_B \circ \alpha$, de manera que f_A es equivalente a f_B.

Recíprocamente, si las funciones lineales f_A y f_B son equivalentes, existen automorfismos $\alpha : \Bbbk^n \to \Bbbk^n$ y $\beta : \Bbbk^m \to \Bbbk^m$ tales que $f_A = \beta \circ f_B \circ \alpha$ y, si $\mathscr{B}$ y $\mathscr{B}'$ son las bases ordenadas estándares de $\Bbbk^n$ y de $\Bbbk^m$, las matrices

$$Q = [\alpha]^{\mathscr{B}}_{\mathscr{B}}, \qquad\qquad P = [\beta]^{\mathscr{B}'}_{\mathscr{B}'}$$

son inversibles y se tiene que

$$A = [f_A]^{\mathscr{B}}_{\mathscr{B}'} = [\beta \circ f_B \circ \alpha]^{\mathscr{B}}_{\mathscr{B}'} = [\beta]^{\mathscr{B}'}_{\mathscr{B}'} \cdot [f_B]^{\mathscr{B}}_{\mathscr{B}'} \cdot [\alpha]^{\mathscr{B}}_{\mathscr{B}} = PBQ.$$

Esto nos dice que A y B son matrices equivalentes. □

6.1.5. El resultado correspondiente a la Proposición 6.1.3 para matrices es:

Proposición. *Sean m, $n \in \mathbb{N}$.*

(*i*) *Dos matrices de $\mathsf{M}_{m,n}(\Bbbk)$ son equivalentes si y solamente si tienen el mismo rango.*

(*ii*) *El número de clases de equivalencia en $\mathsf{M}_{m,n}(\Bbbk)$ es $\min\{m, n\} + 1$.*

(*iii*) *Toda matriz de $\mathsf{M}_{m,n}(\Bbbk)$ es equivalente a exactamente una de la forma*

$$\begin{array}{c c} & \begin{array}{cc} \overbrace{}^{r} & \overbrace{}^{n-r} \end{array} \\ \begin{array}{c} r \\ m-r \end{array} & \begin{pmatrix} I_r & 0 \\ 0 & 0 \end{pmatrix} \end{array} \tag{2}$$

con $r \in [\![0, \min\{m, n\}]\!]$ y dos matrices de éstas son equivalentes solamente si son iguales.

Demostración. Sean A, $B \in \mathsf{M}_{m,n}(\Bbbk)$ y sean $f_A : x \in \Bbbk^n \mapsto Ax \in \Bbbk^m$ y $f_B : x \in \Bbbk^n \mapsto Bx \in \Bbbk^m$. Sabemos del Lema 6.1.4 que las matrices A y B son equivalentes si y solamente si las funciones f_A y f_B lo son, y esto ocurre, según la Proposición 6.1.3*(ii)*, exactamente cuando f_A y f_B tienen el mismo rango. Como A y f_A, por un lado, y B y f_B, por otro, tienen el mismo rango, esto prueba (*i*). Por otro lado, para ver (*ii*) basta observar que el rango de una matriz de $\mathsf{M}_{m,n}(\Bbbk)$ es un elemento de $[\![0, \min\{m, n\}]\!]$ y cada uno los $\min\{m, n\} + 1$ posibles valores ocurre efectivamente como el rango de una matriz de $\mathsf{M}_{m,n}(\Bbbk)$.

Sea, finalmente, $A \in \mathsf{M}_{m,n}(\Bbbk)$ y sea $f_A : x \in \Bbbk^n \mapsto Ax \in \Bbbk^m$. De la Proposición 6.1.3*(i)* sabemos que existen bases ordenadas $\mathscr{B}_1$ y $\mathscr{B}_1'$ de $\Bbbk^n$ y de $\Bbbk^m$ tal que la matriz $[f_A]^{\mathscr{B}_1}_{\mathscr{B}_1'}$ es precisamente la matriz (2) del enunciado. Si $\mathscr{B}_2$ y $\mathscr{B}_2'$ son las bases ordenadas estándares de $\Bbbk^n$ y $\Bbbk^m$, se tiene entonces que

$$A = [f_A]^{\mathscr{B}_2}_{\mathscr{B}_2'} = C(\mathscr{B}_1', \mathscr{B}_2') \cdot [f_A]^{\mathscr{B}_1}_{\mathscr{B}_1'} \cdot C(\mathscr{B}_2, \mathscr{B}_1)$$

y, como las matrices de cambio de base $C(\mathscr{B}_1', \mathscr{B}_2')$ y $C(\mathscr{B}_2, \mathscr{B}_1)$ pertenecen a $\mathsf{GL}_m(\Bbbk)$ y a $\mathsf{GL}_n(\Bbbk)$, esto muestra que A es equivalente a la matriz $[f_A]^{\mathscr{B}_1}_{\mathscr{B}_1'}$, esto es, a la matriz (2) del enunciado. Para terminar, basta observar que esta última matriz tiene rango r, así que si es equivalente a otra de la misma forma necesariamente tienen que ser exactamente iguales. □

6.1.6. Fijemos $n \in \mathbb{N}$.

- Si $i, j \in [\![n]\!]$, escribimos $E^n_{i,j}$ a la matriz de $\mathsf{M}_n(\Bbbk)$ cuya única entrada no nula es la (i, j)-ésima, que es igual a 1.
- Si $i, j \in [\![n]\!]$ son tales que $i \neq j$, escribimos $T^n_{i,j}$ a la matriz de permutación de $\mathsf{M}_n(\Bbbk)$ correspondiente a la transposición $(i\ j)$. Claramente, $T^n_{i,j}$ es inversible y, de hecho, $(T^n_{i,j})^{-1} = T_{i,j}$.
- Si $i, j \in [\![n]\!]$ son tales que $i \neq j$ y $\lambda \in \Bbbk$ es un escalar, ponemos

  $$L^n_{i,j}(\lambda) = I_n + \lambda E^n_{i,j}.$$

  Se trata de una matriz inversible y es fácil verificar que $(L^n_{i,j}(\lambda))^{-1} = L^n_{i,j}(-\lambda)$.
- Si $i \in [\![n]\!]$ y $\lambda \in \Bbbk \setminus 0$, escribimos

  $$D^n_i(\lambda) = I_n + (\lambda - 1)E^n_{i,i}.$$

  Otra vez, esta matriz es inversible y $(D^n_i)(\lambda)^{-1} = D^n_i(\lambda^{-1})$.

Llamamos a las matrices

$$\begin{array}{ll} T^n_{i,j} & \text{con } i, j \in [\![n]\!] \text{ tales que } i \neq j, \\ L^n_{i,j}(\lambda) & \text{con } i, j \in [\![n]\!] \text{ y } \lambda \in \Bbbk \text{ tales que } i \neq j, \text{ y} \\ D^n_i(\lambda) & \text{con } i \in [\![n]\!] \text{ y } \lambda \in \Bbbk \setminus 0 \end{array}$$

las ***matrices elementales*** de $\mathsf{M}_n(\Bbbk)$. Se sigue de lo anterior que todas las matrices elementales son inversibles y que sus inversas son también matrices elementales. Como consecuencia inmediata de esto, vemos que

> *todo producto de matrices elementales es inversible y su matriz inversa es también igual a un producto de matrices elementales.*

6.1.7. Si $m, n \in \mathbb{N}$ y $A \in \mathsf{M}_{m,n}(\Bbbk)$, entonces calculando directamente es fácil ver que:

- Si $i, j \in [\![n]\!]$ e $i \neq j$, la matriz $A' = AT^n_{i,j}$ es la que se obtiene de A intercambiando la columna i-ésima y la j-esima.
- Si $i, j \in [\![n]\!]$ y $\lambda \in \Bbbk$ son tales que $i \neq j$, la matriz $A' = AL^n_{i,j}(\lambda)$ es la que se obtiene de A sumándole a su columna j-ésima el resultado de multiplicar a su columna i-ésima por λ.
- Si $i \in [\![n]\!]$ y $\lambda \in \Bbbk \smallsetminus 0$, la matriz $A' = AD^n_i(\lambda)$ es la que se obtiene de A multiplicando su columna i-ésima por λ.

Decimos en cada uno de estos casos que la matriz A' se obtiene de A haciendo una ***operación de columnas***. De manera similar:

- Si $i, j \in [\![m]\!]$ e $i \neq j$, la matriz $A'' = T^m_{i,j}A$ es la que se obtiene de A intercambiando la fila i-ésima y la j-esima.
- Si $i, j \in [\![n]\!]$ y $\lambda \in \Bbbk$ son tales que $i \neq j$, la matriz $A'' = L^m_{i,j}(\lambda)A$ es la que se obtiene de A sumándole a su fila i-ésima el resultado de multiplicar a su fila j-ésima por λ.
- Si $i \in [\![m]\!]$ y $\lambda \in \Bbbk \smallsetminus 0$, la matriz $A'' = D^m_i(\lambda)A$ es la que se obtiene de A multiplicando su fila i-ésima por λ.

En cada uno de estos casos decimos que la matriz A'' se obtiene de A haciendo una ***operación de filas***.

6.1.8. Una de las razones por las que las matrices elementales son importantes es la siguiente:

Proposición. *Sea $n \in \mathbb{N}$. Toda matriz de $\mathsf{GL}_n(\Bbbk)$ es igual a un producto de matrices elementales.*

Demostración. Sea $(e_1, \dots, e_n)$ la base ordenada estándar de $\Bbbk^n$. Mostremos que

> *existen matrices $P_0, \dots, P_n$, cada una de ellas igual a un producto de matrices elementales, tales que para cada $i \in [\![0, n]\!]$ las primeras i columnas del producto $P_i \cdots P_0 A$ son exactamente los vectores $e_1, \dots, e_i$, en orden.* (3)

Esto probará la proposición ya que entonces tendremos que $P_n \cdots P_0 A = I_n$, ya que la única matriz de $\mathsf{GL}_n(\Bbbk)$ que tiene sus n columnas iguales a los vectores $e_1, \dots, e_n$ en orden es la matriz identidad I_n y, por lo tanto, $A = P_0^{-1} \cdots P_n^{-1}$ es un producto de matrices elementales.

Construimos las matrices $P_0, \dots, P_n$ inductivamente, empezando con la observación de que podemos tomar simplemente $P_0 = I_n$. Supongamos entonces que $r \in [\![0, n-1]\!]$ y que ya construimos matrices $P_0, \dots, P_r$, cada una de ellas igual a un producto de matrices elementales, de manera que el producto $B = P_r \cdots P_0 A$ tiene sus primeras r columnas iguales a los vectores $e_1, \dots, e_r$, en orden. La matriz B tiene entonces la forma

$$\begin{pmatrix}
1 & & & b_{1,r+1} & b_{1,r+2} & \cdots & b_{1,n} \\
& \ddots & & \vdots & \vdots & \ddots & \vdots \\
& & 1 & b_{r,r+1} & b_{r,r+2} & \cdots & b_{r,n} \\
0 & \cdots & 0 & b_{r+1,r+1} & b_{r+1,r+2} & \cdots & b_{r+1,n} \\
0 & \cdots & 0 & b_{r+2,r+1} & b_{r+2,r+2} & \cdots & b_{r+2,n} \\
\vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\
0 & \cdots & 0 & b_{n,r+1} & b_{n,r+2} & \cdots & b_{n,n}
\end{pmatrix}$$

Como B es inversible, sus columnas son elementos linealmente independientes de $\Bbbk^n$. Tiene que existir entonces $s \in [\![r+1, n]\!]$ tal que $b_{s,r+1} \neq 0$: si no fuese ése el caso, la columna $(r+1)$-ésima de B sería una combinación lineal de las primeras r. Si $s = r+1$, pongamos $C_0 = D_s^n(b_{r+1,r+1}^{-1})$ y si $s \neq r+1$ pongamos $C_0 = D_s^n(b_{s,r+1}^{-1})T_{s,r+1}$. En cualquier caso, tenemos que la matriz

$$B' = C_0 B$$

es de la forma

$$\begin{pmatrix} 1 & & & b'_{1,r+1} & b'_{1,r+2} & \cdots & b'_{1,n} \\ & \ddots & & \vdots & \vdots & \ddots & \vdots \\ & & 1 & b'_{r,r+1} & b'_{r,r+2} & \cdots & b'_{r,n} \\ 0 & \cdots & 0 & 1 & b'_{r+1,r+2} & \cdots & b'_{r+1,n} \\ 0 & \cdots & 0 & b'_{r+2,r+1} & b'_{r+2,r+2} & \cdots & b'_{r+2,n} \\ \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & \cdots & 0 & b'_{n,r} & b'_{n,r+2} & \cdots & b'_{n,n} \end{pmatrix}$$

con un 1 en la entrada $(r+1, r+1)$-ésima. Consideremos ahora la matriz B'' que se obtiene de B reemplazando, para cada $i \in [\![1, n]\!] \setminus \{r+1\}$, la fila i-ésima por el resultado de restarle a esa fila la fila $(r+1)$-ésima multiplicada por $b_{i,r+1}$: de acuerdo a nuestras observaciones de 6.1.7, esto significa que

$$B'' = L_{n,r+1}^n(-b_{n,r+1}) \cdots \widehat{L_{n,r+1}^n(-b_{n,r+1})} \cdots L_{r+2,r+1}^n(-b_{r+2,r+1})B'$$

La matriz B'' es de la forma

$$\begin{pmatrix} 1 & & & 0 & b''_{1,r+2} & \cdots & b''_{1,n} \\ & \ddots & & \vdots & \vdots & \ddots & \vdots \\ & & 1 & 0 & b''_{r,r+2} & \cdots & b''_{r,n} \\ 0 & \cdots & 0 & 1 & b''_{r+1,r+2} & \cdots & b''_{r+1,n} \\ 0 & \cdots & 0 & 0 & b''_{r+2,r+2} & \cdots & b''_{r+1,n} \\ \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & \cdots & 0 & 0 & b''_{n,r+2} & \cdots & b''_{n,n} \end{pmatrix}$$

y tiene sus primeras $r+1$ columnas iguales a los vectores $e_1, \dots, e_{r+1}$. Esto significa que si ponemos

$$P_{r+1} = \begin{cases} L_{n,r+1}^n(-b_{n,r+1}) \cdots \widehat{L_{n,r+1}^n(-b_{n,r+1})} \cdots L_{r+2,r+1}^n(-b_{r+2,r+1}) D_s^n(b_{s,r+1}^{-1}), & \text{si } s = r+1; \\ L_{n,r+1}^n(-b_{n,r+1}) \cdots \widehat{L_{n,r+1}^n(-b_{n,r+1})} \cdots L_{r+2,r+1}^n(-b_{r+2,r+1}) D_s^n(b_{s,r+1}^{-1}) T_{s,r+1}, & \text{si } s \neq r+1; \end{cases}$$

entonces $P_{r+1} \cdots P_0 A$ tiene sus primeras $r+1$ columnas iguales a $e_1, \dots, e_{r+1}$. Como manifiestamente P_{r+1} es un producto de matrices elementales, esto prueba la afirmación (3). □

6.1.9. Proposición. *Sean $m, n \in \mathbb{N}$ y $A, B \in \mathsf{M}_{m,n}(\Bbbk)$. Las siguientes afirmaciones son equivalentes:*

(a) *Las matrices A y B son equivalentes.*

(b) *Es posible obtener la matriz A a partir de B haciendo operaciones de filas y columnas.*

Demostración. **Hacer.** □

§2. Conjugación de endomorfismos y de matrices cuadradas

6.2.1. Cuando trabajamos con endomorfismos de un espacio vectorial, es de interés una noción diferente de equivalencia: si V es un espacio vectorial y $f, g : V \to V$ son endomorfismos de V, decimos que f y g son ***conjugados*** si existe un automorfismo $\alpha : V \to B$ tal que $f = \alpha^{-1} \circ g \circ \alpha$, y en ese caso escribimos $f \approx g$. Esta es una relación de equivalencia en el conjunto $\mathsf{End}(V)$:

- Si $f \in \mathsf{End}(V)$, entonces $f = \mathsf{id}_V^{-1} \circ f \circ \mathsf{id}_V$, así que $f \approx f$.
- Si $f, g \in \mathsf{End}(V)$ son tales que $f \approx g$, entonces existe un automorfismo $\alpha : V \to V$ tal que $f = \alpha^{-1} \circ g \circ \alpha$ y se sigue de esto que $g = (\alpha^{-1})^{-1} \circ f \circ \alpha^{-1}$ y, como α^{-1} es un automorfismo de V, que $g \approx f$.
- Si $f, g, h \in \mathsf{End}(V)$ son tales que $f \approx g$ y $g \approx h$, existen automorfismos $\alpha, \beta : V \to V$ tales que $f = \alpha^{-1} \circ g \circ \alpha$ y $g = \beta^{-1} \circ h \circ \beta$, y entonces

$$f = \alpha^{-1} \circ \beta^{-1} \circ h \circ \beta \circ \alpha = (\beta \circ \alpha)^{-1} \circ h \circ (\beta \circ \alpha).$$

  Como la composición $\beta \circ \alpha$ es un automorfismo de V, se sigue de esto que $f \approx h$.

De manera similar, si $n \in \mathbb{N}$, decimos que dos matrices A y B de $\mathsf{M}_n(\Bbbk)$ son ***conjugadas***, y escribimos $A \approx B$, si existe una matriz inversible $P \in \mathsf{GL}_n(\Bbbk)$ tal que $A = P^{-1}BP$. Como para los endomorfismos, la relación de conjugación es una relación de equivalencia sobre $\mathsf{End}(V)$.

6.2.2. Lema. *Sea $n \in \mathbb{N}$ y sean A, $B \in \mathsf{M}_n(\Bbbk)$. Las siguientes afirmaciones son equivalentes:*

(a) *Las matrices A y B son conjugadas.*

(b) *Los endomorfismos $f_A : x \in \Bbbk^n \mapsto Ax \in \Bbbk^n$ y $f_B : x \in \Bbbk^n \mapsto Bx \in \Bbbk^n$ de $\Bbbk^n$ son conjugados.*

Demostración. Si A y B son conjugadas, existe una matriz inversible $P \in \mathsf{GL}_n(\Bbbk)$ tal que $A = P^{-1}BP$ y entonces la función $f_P : x \in \Bbbk^n \mapsto Px \in \Bbbk^n$ es un automorfismo y $f_A = f_P^{-1} \circ f_B \circ f_P$, de manera que los isomorfismos f_A y f_B son conjugados.

Recíprocamente, si f_A y f_B son conjugados, de manera que existe un automorfismo $\alpha : V \to V$ tal que $f_A = \alpha^{-1} \circ f_B \circ \alpha$, y $\mathscr{B}$ es la base ordenada estándar de $\Bbbk^n$, tenemos que la matriz $[\alpha]_{\mathscr{B}}^{\mathscr{B}}$ es inversible, que

$$A = [f_A]_{\mathscr{B}}^{\mathscr{B}} = [\alpha^{-1} \circ f_B \circ \alpha]_{\mathscr{B}}^{\mathscr{B}} = [\alpha^{-1}]_{\mathscr{B}}^{\mathscr{B}} \cdot [f_B]_{\mathscr{B}}^{\mathscr{B}} \cdot [\alpha]_{\mathscr{B}}^{\mathscr{B}} = ([\alpha]_{\mathscr{B}}^{\mathscr{B}})^{-1} \cdot B \cdot [\alpha]_{\mathscr{B}}^{\mathscr{B}}$$

y, por lo tanto, que las matrices A y B son conjugadas. □

6.2.3. Como en el caso de la equivalencia de funciones lineales, la relación de conjugación entre endomorfismos tiene una interpretación concreta en términos de matrices muy sencilla, análoga a la de la Proposición 6.1.2:

Proposición. *Sea V un espacio vectorial de dimensión finita. Dos endomorfismos $f, g : V \to V$ son conjugados si y solamente si existen bases ordenadas $\mathscr{B}$ y $\mathscr{B}'$ de V tales que $[f]^{\mathscr{B}}_{\mathscr{B}} = [g]^{\mathscr{B}'}_{\mathscr{B}'}$.*

Demostración. HACER. □

6.2.4. La descripción de las clases de equivalencia para la relación de conjugación en $\mathsf{End}(V)$ y en $\mathsf{M}_n(\mathbb{k})$ es considerablemente más difícil que para la de equivalencia y ocupará el resto de este capítulo: nuestro objetivo es obtener un resultado como los de las Proposiciones 6.1.3 y 6.1.5.

6.2.5. Es importante notar que la relación de conjugación es estrictamente más restrictiva que la de equivalente. Dos endomorfismos $f, g : V \to V$ de un espacio vectorial V son equivalentes si son conjugados, pero no vale la implicación recíproca. Por ejemplo, las matrices $A = \left(\begin{smallmatrix} 0 & 0 \\ 1 & 0 \end{smallmatrix}\right)$ y $B = \left(\begin{smallmatrix} 1 & 0 \\ 0 & 0 \end{smallmatrix}\right)$ son equivalentes, porque tienen claramente el mismo rango, pero no son conjugadas. En efecto, sabemos de la Proposición 5.3.4 que si lo fueran ambas tendrían el mismo polinomio característico, pero $\chi_A = X^2$ mientras que $\chi_B = X(X-1)$.

§3. Endomorfismos descomponibles e indescomponibles

6.3.1. Decimos que un endomorfismo $f : V \to V$ de un espacio vectorial V no nulo es ***descomponible*** si existen subespacios no nulos y f-invariantes W_1 y W_2 de V tales que $V = W_1 \oplus W_2$, y que es ***indescomponible*** si no es descomponible y $V \neq 0$. Observemos que, de acuerdo a estas definiciones, el endomorfismo nulo del espacio vectorial nulo $0 : 0 \to 0$ no es ni descomponible ni indescomponible.

La descomponibilidad tiene una interpretación sencilla en términos de matrices:

Proposición. *Un endomorfismo $f : V \to V$ de un espacio vectorial V de dimensión finita y positiva es descomponible si y solamente si existe una base $\mathscr{B}$ de V tal que la matriz de f con respecto a $\mathscr{B}$ es diagonal por bloques de manera no trivial.*

Demostración. Si $f : V \to V$ es un endomorfismo descomponible de un espacio vectorial de dimensión finita y $V = V_1 \oplus V_2$ es una descomposición en suma directa de V como suma de subespacios no nulos y f-invariantes, cada vez que $\mathscr{B}_1 = (x_1, \dots, x_r)$ y $\mathscr{B}_2 = (y_1, \dots, y_s)$ son bases ordenadas de V_1 y V_2, respectivamente, se tiene que $\mathscr{B} = (x_1, \dots, x_r, y_1, \dots, y_s)$ es una base ordenada de V y que la matriz de f con respecto a $\mathscr{B}$ es diagonal por bloques de manera no trivial: si $f_{V_1} : V_1 \to V_1$ y $f_{V_2} : V_2 \to V_2$ son las restricciones de f a V_1 y a V_2, es claro que

$$[f]^{\mathscr{B}}_{\mathscr{B}} = \begin{pmatrix} [f_{V_1}]^{\mathscr{B}_1}_{\mathscr{B}_1} & 0 \\ 0 & [f_{V_2}]^{\mathscr{B}_2}_{\mathscr{B}_2} \end{pmatrix},$$

Recíprocamente, si $f : V \to V$ es un endomorfismo de un espacio vectorial de dimensión finita n tal que existe una base ordenada $\mathscr{B} = (z_1, \dots, z_n)$ de V tal que la matriz $[f]^{\mathscr{B}}_{\mathscr{B}}$ es diagonal por bloques,

$$[f]^{\mathscr{B}}_{\mathscr{B}} = \begin{pmatrix} A_1 & 0 \\ 0 & A_1 \end{pmatrix}$$

con $A_1 \in \mathsf{M}_r(\Bbbk)$ y $A_2 \in \mathsf{M}_s(\Bbbk)$, entonces por supuesto $n = r + s$ y los subespacios $V_1 = \langle z_1, \dots, z_r \rangle$ y $V_2 = \langle z_{r+1}, \dots, z_n \rangle$ son no nulos, f-invariantes y tales que $V = V_1 \oplus V_2$, de manera que el endomorfismo f es descomponible. □

6.3.2. Ejemplo. HACER. ◇

6.3.3. El siguiente resultado nos dice que siempre podemos descomponer un endomorfismo de un espacio vectorial de dimensión finita en "componentes indescomponibles":

Proposición. *Sea V un espacio vectorial de dimensión finita positiva y sea $f : V \to V$ un endomorfismo. Existen $n \geq 1$ y subespacios f-invariantes $V_1, \dots, V_n$ de V tales que*

- *para cada $i \in [\![n]\!]$ la restricción $f_{V_i} : V_i \to V_i$ es indescomponible, y*
- *$V = V_1 \oplus \cdots \oplus V_n$.*

Demostración. Existen descomposiciones $V = W_1 \oplus \cdots \oplus W_r$ de V como suma directa de subespacios no nulos f-invariantes: por ejemplo, podemos tomar $r = 1$ y $W_1 = V$. Como el número de sumandos en una tal descomposición es claramente a lo sumo igual a $\dim V$, podemos considerar el número máximo n de sumandos de una tal descomposición de V y fijar una descomposición $V = V_1 \oplus \cdots \oplus V_n$ de V como suma directa de exactamente n subespacios no nulos f-invariantes.

Ahora bien, si existiese $i \in [\![n]\!]$ tal que la restricción f_{V_i} fuera descomponible, habría subespacios no nulos y f-invariantes U_1 y U_2 de V_i tales que $V_i = U_1 \oplus U_2$ y, en consecuencia,

$$V = V_1 \oplus \cdots \oplus V_{i-1} \oplus U_1 \oplus U_2 \oplus V_{i+1} \oplus \cdots \oplus V_n$$

sería una descomposición de V como suma directa de subespacios no nulos y f-invariantes con $n + 1$ sumandos, lo que es imposible. Esta significa que cada una de las restricciones $f_{V_1}, \dots, f_{V_n}$ es indescomponible y prueba la proposición. □

6.3.4. Es importante observar que en general la descomposición de un espacio vectorial V en componentes indescomponibles con respecto a un endomorfismo $f : V \to V$ que nos da la Proposición 6.3.3 no es única. Por ejemplo, si $f = \mathrm{id}_V$ es la función identidad de V y $\dim V > 1$, entonces cualquier descomposición de V como suma directa de subespacios de dimensión 1 satisface esas condiciones.

6.3.5. Una de las razones por las que los endomorfismos indescomponibles son importantes es que tienen polinomios minimales de una forma especial. Este punto será central varias veces en lo que sigue.

Proposición. *Sea V un espacio vectorial de dimensión finita y positiva. Si $f : V \to V$ es un endomorfismo indescomponible de V, entonces existen un polinomio mónico e irreducible $p \in \Bbbk[X]$ y un entero positivo ν tales que el polinomio minimal de f es $m_f = p^\nu$.*

Demostración. Sabemos que existen un entero positivo r, polinomios mónicos, irreducibles y coprimos dos a dos $p_1, \dots, p_r$ y enteros positivos $v_1, \dots, v_r$ tales que $m_f = p_1^{v_1} \cdots p_r^{v_r}$, y la Proposición 5.7.1 nos dice que los subespacios $V_i = \mathsf{Nu}((p_i^{v_i})(f))$ con $i \in [\![n]\!]$ son no nulos, f-invariantes y tales que $V = V_1 \oplus \cdots \oplus V_n$. No puede ser que n sea mayor que 1, porque f es indescomponible, y esto prueba la proposición. □

§4. Endomorfismos nilpotentes

6.4.1. Si V es un espacio vectorial, un endomorfismo $f : V \to V$ de V es ***nilpotente*** si existe $k \in \mathbb{N}_0$ tal que $f^k = 0$ y en ese caso el número $\min\{i \in \mathbb{N}_0 : f^i = 0\}$ es el ***índice de nilpotencia*** de f.

Proposición. *Sea V un espacio vectorial de dimensión finita n y sea $f : V \to V$ un endomorfismo de V. Las siguientes afirmaciones son equivalentes:*

(a) *f es nilpotente.*
(b) $f^n = 0$.
(c) *Existe $k \in \mathbb{N}_0$ tal que el polinomio minimal de f es $m_f = X^k$.*
(d) *El polinomio característico de f es $\chi_f = X^n$.*

Si se cumplen, el entero k que aparece en (c) *es el índice de nilpotencia de f y es a lo sumo igual a n.*

Demostración. Para cada $i \in \mathbb{N}_0$ sea $V_i = \mathsf{Nu}(f^i)$. Es inmediato verificar que $V_i \subseteq V_{i+1}$ para todo $i \in \mathbb{N}_0$, así que tenemos una cadena creciente de subespacios de V,

$$V_0 \subseteq V_1 \subseteq V_2 \subseteq \cdots \tag{4}$$

La sucesión de las dimensiones de estos subespacios

$$\dim V_0 \leq \dim V_1 \leq \dim V_1 \leq \cdots$$

es no decreciente y está acotada por $\dim V$, así que necesariamente se estabiliza: existe $i_0 \in \mathbb{N}_0$ tal que para cada $i \geq i_0$ es $\dim V_i = \dim V_{i_0}$ y, por lo tanto, $V_i = V_{i_0}$, ya que como $V_{i_0} \subseteq V_i$. Podemos suponer además que elegimos i_0 lo más chico posible, y esto implica que cada una de las primeras i_0 inclusiones de la cadena (4)

$$V_0 \subsetneq V_1 \subsetneq \cdots \subsetneq V_{i_0}$$

es propia y, en consecuencia, que

$$\dim V_0 < \dim V_1 < \cdots < \dim V_{i_0}.$$

De esto y de que $V_0 = 0$ se deduce que

$$\dim V_{i_0} = \dim V_{i_0} - \dim V_0 = \sum_{i=1}^{i_0} (\dim V_i - \dim V_{i-1}) \geq i_0,$$

ya que uno de los sumandos de la suma vale por lo menos 1. Probemos ahora la proposición.

$(a) \Rightarrow (b)$ Como f es nilpotente, existe $k \in \mathbb{N}_0$ tal que $f^k = 0$ y entonces $V_k = V$. Esto implica que $V_i = V$ para todo $i \geq k$. Como por otro lado es $V_i = V_{i_0}$ si $i \geq i_0$, se sigue de esto que $V_{i_0} = V$, de manera que por un lado $f^{i_0} = 0$ y, por otro, que $\dim V_{i_0} = n$, con lo que $n \geq i_0$. De estas dos cosas deducimos que $f^n = f^{n-i_0} f^{i_0} = 0$, como afirma (b).

$(b) \Rightarrow (c)$ Si ponemos $p = X^n \in \mathbb{k}[X]$, entonces la hipótesis es que $f(f) = 0$ y la propiedad característica del polinomio minimal m_f implica que m_f divide a X^n: esto significa, claro, que existe $k \in [\![0, n]\!]$ tal que $m_f = X^k$.

$(c) \Rightarrow (d)$ De la Proposición 5.7.4 sabemos que el polinomio minimal m_f y el polinomio característico χ_f de f tienen los mismos divisores irreducibles. Si $m_f = X^k$ para algún $k \in \mathbb{N}_0$, entonces el único divisor irreducible de m_f es X, así que χ_f tiene que ser de la forma X^l para algún $l \in \mathbb{N}_0$. Como el grado de χ_f es igual a la dimensión de V, se tiene entonces que $\chi_f = X^n$.

$(d) \Rightarrow (a)$ Si el polinomio característico de f es $\chi_f = X^n$, entonces el Teorema de Cayley–Hamilton 5.6.9 nos dice que $f^n = \chi_f(f) = 0$: el endomorfismo f es por lo tanto nilpotente.

Esto completa la prueba de la equivalencia de las cuatro afirmaciones del enunciado. Si suponemos que valen, de manera que existe $k \in \mathbb{N}_0$ tal que $m_f = X^k$, entonces $f^k = m_f(f) = 0$ y $f^{k-1} \neq 0$ porque f no anula al polinomio X^{k-1}, ya que éste tiene grado menor que el de m_f. Esto significa que k es el índice de nilpotencia de f. □

6.4.2. Ejemplos.

(a) Si $f : V \to V$ es un endomorfismo nilpotente de un espacio vectorial V y W es un subespacio f-invariante de V, entonces la restricción $f_W : W \to W$ es nilpotente y su índice de nilpotencia no es mayor que el de f.

(b) Si $n \in \mathbb{N}$ y $A \in \mathsf{M}_n(\mathbb{k})$ es una matriz triangular inferior, entonces A es nilpotente si y solamente si todas las entradas de su diagonal son nulas. En efecto, si $A = (a_{i,j})$, sabemos que el polinomio característico de A es $\chi_A = (X - a_{1,1}) \cdots (X - a_{n,n})$ y A es nilpotente, de acuerdo a la Proposición 6.4.1, exactamente cuando este polinomio es igual a X^n. **Hacer:** Definir arriba nilpotencia para matrices

(c) Un endomorfismo $f : V \to V$ diagonalizable es nilpotente si y solamente si es nulo. En efecto, como es diagonalizable existe una base $\mathscr{B}$ de V cuyos elementos son autovectores de f. Para cada $x \in \mathscr{B}$ sea λ_x el autovalor de f correspondiente a x, de manera que $f(x) = \lambda_x x$.

Supongamos que f es nilpotente y sea $k \in \mathbb{N}_0$ tal que $f^k = 0$. Si $x \in \mathscr{B}$, entonces $0 = f^k(x) = \lambda_x^k x$, de manera que $\lambda_x = 0$ y, en particular, $f(x) = \lambda_x x = 0$. Vemos así que f se anula en cada uno de los elementos de la base $\mathscr{B}$ y que es, por lo tanto, el endomorfismo nulo de V. Recíprocamente, es claro que si f es nulo, entonces es nilpotente.

(d) Sea $n \in \mathbb{N}_0$. El endomorfismo $D : p \in \mathbb{R}[X]_{\leq n} \mapsto p' \in \mathbb{R}[X]_{\leq n}$ es en endomorfismo del espacio vectorial $\mathbb{R}[X]_{\leq n}$ de polinomios con coeficientes reales de grado a lo sumo n dado por la derivación es nilpotente. En efecto, sabemos que $D^{n+1}(p) = 0$ si p es un polinomio nulo o de grado a lo sumo n. Más aún, el índice de nilpotencia de D es precisamente $n+1$: sabemos que ese índice está acotado por $\dim \mathbb{R}[X]_{\leq n} = n+1$ y, por otro lado, la potencia

n-ésima de D no es nula, ya que $D^n(X^n) = n!$.

De manera similar, es fácil ver que el endomorfismo

$$L : p \in \mathbb{R}[X]_{\leq n} \mapsto p(X) - p(X-1) \in \mathbb{R}[X]_{\leq n}$$

es nilpotente de índice de nilpotencia igual a $n+1$.

(e) Sean ahora p un número primo y n un entero no negativo, y consideremos la función lineal $D : p \in \mathbb{F}_p[X]_{\leq n} \mapsto p' \in \mathbb{F}_p[X]_{\leq n}$ dada por derivación formal de polinomios con coeficientes en el cuerpo $\mathbb{F}_p$. Se trata del único endomorfismo de $\mathbb{F}_p[X]_{\leq n}$ tal que $D(X^i) = iX^{i-1}$ para todo $i \in [\![0, n]\!]$. Una inducción evidente muestra que para cada $i \in [\![0, n]\!]$ vale que

$$D^p(X^i) = i(i-1)\cdots(i-p+1)X^{i-p}$$

y —como uno de los p enteros $i, i-1, \ldots, i-p+1$ es divisible por p y, por lo tanto, igual a 0 en $\mathbb{F}_p$— esto implica que $D^p = 0$. Así, D es un endomorfismo nilpotente de $\mathbb{F}_p[X]_{\leq n}$ y su índice de nilpotencia es $\min\{n, p\}$. ◇

6.4.3. El resultado más importante sobre endomorfismos nilpotentes es que cuando son indescomponibles admiten un vector cíclico:

Proposición. *Sea V un espacio vectorial de dimensión finita y positiva n. Si $f : V \to V$ es un endomorfismo nilpotente e indescomponible, entonces existe $x \in V$ tal que $V = \langle x \rangle_f$, el índice de nilpotencia de f es igual a la dimensión de V y existe una base ordenada $\mathscr{B}$ de V tal que la matriz de f con respecto a $\mathscr{B}$ es*

$$N_n = \begin{pmatrix} 0 & & & & \\ 1 & 0 & & & \\ & 1 & 0 & & \\ & & \ddots & \ddots & \\ & & & 1 & 0 \end{pmatrix}$$

Llamamos a la matriz N_n que aparece aquí un ***bloque de Jordan nilpotente de tamaño n***.

Demostración. Procedemos por inducción con respecto a la dimensión de V, observando que si $\dim V = 1$ no hay nada que probar.

Supongamos entonces que $\dim V > 1$ y que $f : V \to V$ es un endomorfismo nilpotente e indescomponible. Como f no es sobreyectivo, el subespacio $\mathrm{Im}(f)$ de V es propio y existe un subespacio W de V tal que $\mathrm{Im}(f) \subseteq W \subsetneq V$ y $\dim W = \dim V - 1$. Este subespacio es f-invariante, ya que $f(W) \subseteq \mathrm{Im}(f) \subseteq W$, y la restricción $f_W : W \to W$ de f a W es nilpotente.

La Proposición 6.3.3 nos dice que hay un entero positivo n y subespacios f-invariantes y no nulos $W_1, \ldots, W_n$ de W tales que $W = W_1 \oplus \cdots \oplus W_n$ y cada restricción $f_{W_i} : W_i \to W_i$ es indescomponible. Como estas restricciones son nilpotentes y los subespacios $W_1, \ldots, W_n$ tienen todos dimensión menor que la de V, podemos suponer inductivamente que existen vectores no nulos $x_1, \ldots, x_n \in V$ tales que $W_i = \langle x_i \rangle_f$ para cada $i \in [\![n]\!]$, de manera que

$$W = \langle x_1 \rangle_f \oplus \cdots \oplus \langle x_n \rangle_f. \tag{5}$$

Por otro lado, como $\dim W = \dim V - 1$ hay un vector no nulo $x \in V$ tal que $V = \langle x \rangle \oplus W$. En definitiva, tenemos una descomposición

$$V = \langle x \rangle \oplus \langle x_1 \rangle_f \oplus \cdots \oplus \langle x_n \rangle_f.$$

Como $f(x) \in \operatorname{Im}(f) \subseteq W = \langle x_1 \rangle_f \oplus \cdots \oplus \langle x_n \rangle_f$, existen polinomios $p_1, \ldots, p_n \in \Bbbk[X]$ tales que

$$f(x) = p_1(f)(x_1) + \cdots + p_n(f)(x_n). \tag{6}$$

Para cada $i \in [\![n]\!]$ hay un escalar $\lambda_i \in \Bbbk$ y un polinomio $q_i \in \Bbbk[X]$ tal que $p_i = \lambda_i + Xq_i$. El vector

$$y = x - q_1(f)(x_1) - \cdots - q_n(f)(x_n),$$

no pertenece a W, así que $V = \langle y \rangle \oplus W$ y, por lo tanto,

$$V = \langle y \rangle \oplus \langle x_1 \rangle_f \oplus \cdots \oplus \langle x_n \rangle_f. \tag{7}$$

Por otro lado, la igualdad (6) y la forma en que elegimos a y implican que

$$\begin{aligned} f(y) &= f(x) - f(q_1(f)(x_1)) - \cdots - f(q_n(f)(x_n)) \\ &= p_1(f)(x_1) + \cdots + p_n(f)(x_n) - (Xq_1)(f)(x_1) - \cdots - (Xq_n)(f)(x_n) \\ &= (p_1 - Xq_1)(f)(x_1) + \cdots + (p_n - Xq_n)(f)(x_n) \\ &= \lambda_1 x_1 + \cdots + \lambda_1 x_n. \end{aligned}$$

Si todos los escalares $\lambda_1, \ldots, \lambda_n$ fuesen nulos, entonces tendríamos que $f(y) = 0$ y todos los sumandos directos de la descomposición (7), que son todos no nulos y por lo menos dos, serían f-invariantes: esto es imposible porque f es indescomponible. Tiene sentido entonces considerar el número

$$r = \max\{\dim\langle x_i \rangle_f : i \in [\![n]\!], \lambda_i \neq 0\}$$

y, de hecho, podemos suponer sin pérdida de generalidad que $\lambda_1 \neq 0$ y que $\dim\langle x_1 \rangle_f = r$.

Sea $y_1 = \lambda_1 x_1 + \cdots + \lambda_n x_n$, de manera que $f(y) = y_1$. El subespacio

$$\langle y_1 \rangle_f + \langle x_2 \rangle_f + \cdots + \langle x_n \rangle_f.$$

de W es f-invariante y contiene a y_1 y a $x_2, \ldots, x_n$, así que como $\lambda_1 \neq 0$, también contiene a x_1. Esto implica, en vista de (5), que coincide con W. Además, como $f^r(x_i) = 0$ para todo $i \in [\![n]\!]$ tal que $\lambda_i \neq 0$ por la forma en que elegimos a r, es claro que $f^r(y_1) = 0$, de manera que $\dim\langle y_1 \rangle_f \leq r$. Se sigue de esto que

$$\begin{aligned} \dim W &\leq \dim\langle y_1 \rangle_f + \dim\langle x_2 \rangle_f + \cdots + \dim\langle x_n \rangle_f \\ &\leq r + \dim\langle x_2 \rangle_f + \cdots + \dim\langle x_n \rangle_f \\ &= \dim\langle x_1 \rangle_f + \dim\langle x_2 \rangle_f + \cdots + \dim\langle x_n \rangle_f = \dim W \end{aligned}$$

y, por lo tanto, que valen las igualdades. Usando la Proposición 1.10.5 podemos concluir entonces que $W = \langle y_1 \rangle_f \oplus \langle x_2 \rangle_f \oplus \cdots \oplus \langle x_n \rangle_f$ y, en consecuencia, que

$$V = \langle y \rangle \oplus \langle y_1 \rangle_f \oplus \langle x_2 \rangle_f \oplus \cdots \oplus \langle x_n \rangle_f.$$

En esta descomposición de V, el sumando $\langle y \rangle \oplus \langle y_1 \rangle_f$ y cada uno de lo sumandos $\langle x_i \rangle_f$ con $i \in [\![2, n]\!]$ son f-invariantes y no nulos. Como f es indescomponible e $y \neq 0$, vemos que debe ser necesariamente $n = 1$, con lo que $V = \langle y \rangle \oplus \langle y_1 \rangle_f = \langle y \rangle_f$.

Esto prueba la primera afirmación del enunciado. La segunda es ahora inmediata y para probar la tercera es suficiente notar que podemos tomar $\mathscr{B} = (x, f(x), \dots, f^{n-1}(x))$. □

6.4.4. Si $f : V \to V$ es un endomorfismo de un espacio vectorial de dimensión finita e $i \in \mathbb{N}_0$, definimos el número $\delta_i(f) = \dim \mathsf{Nu}(f^i)$.

Proposición. *Sea V un espacio vectorial de dimensión finita y positiva n.*

(*i*) *Si $f, g : V \to V$ son endomorfismos conjugados, entonces para todo $i \in \mathbb{N}_0$ es $\delta_i(f) = \delta_i(g)$.*

(*ii*) *Si $f : V \to V$ es un endomorfismo y $V_1, \dots, V_r$ son subespacios f-invariantes de V tales que $V = V_1 \oplus \cdots \oplus V_r$, entonces para todo $i \in \mathbb{N}_0$ se tiene que $\delta_i(f) = \delta_i(f_{V_1}) + \cdots + \delta_i(f_{V_r})$.*

(*iii*) *Si $f : V \to V$ un endomorfismo nilpotente e indescomponible de V, entonces para todo $i \in \mathbb{N}$ se tiene que*

$$-\delta_{i+1}(f) + 2\delta_i(f) - \delta_{i-1}(f) = \begin{cases} 1, & \text{si } i = n; \\ 0, & \text{en caso contrario.} \end{cases}$$

Demostración. (*i*) Sean $f, g : V \to V$ dos endomorfismos conjugados y sea $\alpha : V \to V$ un automorfismo tal que $f = \alpha^{-1} \circ g \circ \alpha$. Notemos que $\delta_0(f) = 0 = \delta_0(g)$.

Si $x \in \mathsf{Nu}(f)$, tenemos que

$$g(\alpha(x)) = \alpha\big((\alpha^{-1} \circ g \circ \alpha)(x)\big) = \alpha(f(x)) = 0$$

y entonces $\alpha(x) \in \mathsf{Nu}(g)$. Como consecuencia de esto, podemos considerar la función

$$\beta : x \in \mathsf{Nu}(f) \mapsto \alpha(x) \in \mathsf{Nu}(g).$$

De manera similar, tenemos una función

$$\gamma : y \in \mathsf{Nu}(g) \mapsto \alpha^{-1}(y) \in \mathsf{Nu}(f)$$

y es inmediato verificar que β y γ son isomorfismos inversos. En particular, vale que

$$\delta_1(f) = \dim \mathsf{Nu}(f) = \dim \mathsf{Nu}(g) = \delta_1(g).$$

Sea ahora $i \geq 2$. Para probar que $\delta_i(f) = \delta_i(g)$ basta observar que f^i y g^i son endomorfismos de V conjugados, ya que

$$f^i = (\alpha^{-1} \circ g \circ \alpha)^i = (\alpha^{-1} \circ g \circ \alpha) \circ (\alpha^{-1} \circ g \circ \alpha) \circ \cdots \circ (\alpha^{-1} \circ g \circ \alpha) = \alpha^{-1} \circ g^i \circ \alpha,$$

y que entonces, en vista de lo que ya probamos para δ_1,

$$\delta_i(f) = \delta_1(f^i) = \delta_1(g^i) = \delta_i(g).$$

(*ii*) Sea $f : V \to V$ un endomorfismo de V, sean $V_1, \dots, V_r$ subespacios f-invariantes de V tales que

$$V = V_1 \oplus \cdots \oplus V_r \tag{8}$$

y sea $i \in \mathbb{N}_0$. Para ver que $\delta_i(f) = \delta_i(f_{V_1}) + \cdots + \delta_i(f_{V_r})$, es decir, que

$$\dim \mathsf{Nu}(f^i) = \dim \mathsf{Nu}(f^i_{V_1}) + \cdots + \dim \mathsf{Nu}(f^i_{V_r})$$

es suficiente con mostrar que

$$\mathsf{Nu}(f^i) = \mathsf{Nu}(f^i_{V_1}) \oplus \cdots \oplus \mathsf{Nu}(f^i_{V_r}). \tag{9}$$

Supongamos que $x \in \mathsf{Nu}(f^i)$. Como tenemos la descomposición (8), existen $x_1 \in V_1, \dots, x_r \in V_r$ tales que $x = x_1 + \cdots + x_r$ y, por lo tanto,

$$0 = f^i(x) = f^i(x_1) + \cdots + f^i(x_r) = f^i_{V_1}(x_1) + \cdots + f^i_{V_r}(x_r).$$

Como la suma (8) es directa y $f^i_{V_j}(x_j) \in V_j$ para cada $j \in [\![r]\!]$, se sigue de esto que para cada $j \in [\![r]\!]$, de hecho, $f_{V_j}(x_j) = 0$, esto es, que $x_j \in \mathsf{Nu}(f^i_{V_j})$. Esto muestra que

$$\mathsf{Nu}(f^i) = \mathsf{Nu}(f^i_{V_1}) + \cdots + \mathsf{Nu}(f^i_{V_r}).$$

Esta igualdad, el hecho evidente de que $\mathsf{Nu}(f^i_{V_j}) \subseteq V_j$ para cada $j \in [\![r]\!]$, y la descomposición directa (8) implican, de acuerdo a la Proposición 1.10.6, que tenemos la descomposición (9) que queremos.

(*iii*) Sea $f : V \to V$ un endomorfismo nilpotente e indescomponible de V. De la Proposición 6.4.3 sabemos que existe una base ordenada $\mathscr{B} = (x_1, \dots, x_n)$ de V tal que

$$[f]^{\mathscr{B}}_{\mathscr{B}} = \begin{pmatrix} 0 & & & & \\ 1 & 0 & & & \\ & 1 & 0 & & \\ & & \ddots & \ddots & \\ & & & 1 & 0 \end{pmatrix}$$

Calculando explícitamente las potencias de esta matriz vemos que para cada $i \geq 0$ es

$$\delta_i(f) = \begin{cases} i, & \text{si } i \leq n; \\ n, & \text{si } i > n; \end{cases}$$

y entonces

$$-\delta_{i+1}(f)+2\delta_i(f)-\delta_{i-1}(f)=\begin{cases}-(i+1)+2i+(i-1), & \text{si } i\leq n-1;\\ -n+2i-(i-1), & \text{si } i=n;\\ -n+2n-(i-1), & \text{si } i=n+1;\\ -n+2n-n, & \text{si } i>n+1;\end{cases}$$
$$=\begin{cases}1, & \text{si } i=n;\\ 0, & \text{en caso contrario.}\end{cases}$$

Esto es lo que afirma el enunciado. □

6.4.5. Proposición. *Sea V un espacio vectorial de dimensión finita y positiva n y sea $f : V \to V$ un endomorfismo nilpotente de V.*

(*i*) *Existe una única secuencia finita y débilmente decreciente $(n_1, n_2, \dots, n_r)$ de enteros positivos con $n_1 + \cdots + n_r = n$ tal que hay una base ordenada $\mathscr{B}$ de V con respecto a la que la matriz de f es diagonal por bloques, de la forma*

$$[f]_{\mathscr{B}}^{\mathscr{B}} = \begin{pmatrix} N_{n_1} & & & \\ & N_{n_2} & & \\ & & \ddots & \\ & & & N_{n_r}\end{pmatrix}$$

con cada N_{n_i} un bloque de Jordan nilpotente de tamaño n_i.

(*ii*) *Si $s \in \mathbb{N}$ y μ_s es el número de bloques de Jordan nilpotentes de tamaño s que aparecen en esta matriz, de manera que $\mu_s = \#\{i \in [\![r]\!] : n_i = s\}$, entonces*

$$\mu_s = -\delta_{s+1}(f) + 2\delta_i(f) - \delta_{i-1}(f). \tag{10}$$

La cantidad total de bloques es igual a $\delta_1(f)$. **Hacer:** *Hacer esto.*

Llamamos a la secuencia $(n_1, \dots, n_r)$ el ***tipo*** del endomorfismo f.

Demostración. De acuerdo a la Proposición 6.3.3, existen $r \in \mathbb{N}$ y subespacios f-invariantes $V_1, \dots, V_r$ de V tales que $V = V_1 \oplus \cdots \oplus V_r$ y para cada $i \in [\![r]\!]$ la restricción $f_{V_i} : V_i \to V_i$ de f a V_i es indescomponible. Sin pérdida de generalidad, podemos suponer que ademas que la secuencia $(n_1, \dots, n_r) = (\dim V_1, \dots, \dim V_r)$ es débilmente decreciente, ya que si no es ése el caso podemos reindexar los subespacios.

Como $f^n = 0$, es claro que $(f_{V_i})^n = 0$ para cada $i \in [\![n]\!]$, así que los endomorfismos $f_{V_1}, \dots, f_{V_r}$ son nilpotentes. Son además indescomponibles, así que la Proposición 6.4.3 nos dice que para cada $i \in [\![r]\!]$ hay un vector $x_i \in V_i$ con $V_i = \langle x_i \rangle_f$, que $\mathscr{B}_i = (x_i, f(x_i), \dots, f^{n_i-1}(x_i))$ es una base ordenada de V_i, y que

$$[f_{V_i}]_{\mathscr{B}_i}^{\mathscr{B}_i} = N_{n_i},$$

un bloque de Jordan nilpotente de tamaño n. Si $\mathscr{B} = \mathscr{B}_1 \cup \cdots \cup \mathscr{B}_r$ es la base ordenada de V que se obtiene concatenando, en orden, las bases $\mathscr{B}_1, \ldots, \mathscr{B}_r$, es claro que la matriz $[f]_{\mathscr{B}}^{\mathscr{B}}$ es la matriz descripta en el enunciado. Esto prueba la existencia que se afirma en la parte (*a*) de la proposición.

De acuerdo a la segunda afirmación de la Proposición 6.4.4, tenemos que

$$\delta_i(f) = \delta_i(f_{V_1}) + \cdots + \delta_i(f_{V_r})$$

para todo $i \in \mathbb{N}_0$ y entonces, si $s \in \mathbb{N}$ se tiene que

$$-\delta_{s+1}(f) + 2\delta_s(f) - \delta_{s-1}(f) = \sum_{i=1}^{r}\bigl(-\delta_{s+1}(f_{V_i}) + 2\delta_s(f_{V_i}) - \delta_{s-1}(f_{V_i})\bigr).$$

Esto, de acuerdo a la tercera parte de aquel lema, es igual a μ_s, el numero $\#\{i \in [\![r]\!] : \dim V_i = s\}$: en efecto, el término de la suma correspondiente a $i \in [\![r]\!]$ es igual a 1 si $\dim V_i = s$ y es igual a 0 en caso contrario, así que la suma cuenta cuantos de los subespacios V_i tienen dimensión s. Esto prueba la segunda afirmación de la proposición.

Ahora bien, la secuencia de números $\mu_0, \mu_1, \mu_2, \ldots$ está unívocamente determinada por f, ya que el lado derecho de la igualdad (10) depende solamente de f, y es claro que esta secuencia determina completamente al tipo $(n_1, \ldots, n_r)$ de f. Esto prueba la afirmación de unicidad hecha en la parte (*i*) de la proposición. □

6.4.6. Si $f : V \to V$ es un endomorfismo nilpotente de un espacio vectorial de dimensión finita y positiva n, entonces el tipo de f es una secuencia finita y débilmente decreciente $(n_1, \ldots, n_r)$ de enteros positivos tal que $n_1 + \cdots + n_r = n$. Decimos que una tal secuencia es una ***partición*** de n. Por ejemplo, las particiones de 7 son las siguientes quince secuencias:

$$\begin{array}{lllll}
(1,1,1,1,1,1,1), & (2,1,1,1,1,1), & (2,2,1,1,1), & (2,2,2,1), & (3,1,1,1,1), \\
(3,2,1,1), & (3,2,2), & (3,3,1), & (4,1,1,1), & (4,2,1), \\
(4,3), & (5,1,1), & (5,2), & (6,1), & (7).
\end{array}$$

En el Cuadro 6.1 está tabulado el número de particiones de los primeros 15 enteros positivos. Ese número crece muy rápidamente y, por ejemplo, hay

$$24,061,467,864,032,622,473,692,149,727,991$$

particiones de 1000. G. H. Hardy (1877–1947, Inglaterra) y Srinivasa Ramanujan (1887–1920, India) probaron en 1918 que el número de particiones de n es aproximadamente igual a

$$\frac{1}{4n\sqrt{3}} \exp\left(\pi\sqrt{\frac{2n}{3}}\right)$$

cuando n es grande y, de hecho, se conocen aproximaciones mucho mejores y hasta fórmulas exactas para ese número. El estudio de las particiones de enteros es parte de la llamada *teoría de números aditiva* y es una bella, importante y profunda parte de la matemática. Una buena introducción al tema es el libro [And98] de George E. Andrews (1938–, Estados Unidos).

n	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	$\cdots$
p_n	1	2	3	5	7	11	15	22	30	42	56	77	101	135	176	$\cdots$

Cuadro 6.1. El número p_n de particiones de un entero positivo n.

6.4.7. La razón por las que nos interesa el tipo de un endomorfismo nilpotente es que nos permite resolver el problema de decidir si dos endomorfismos son conjugados o no:

Proposición. *Sea V un espacio vectorial de dimensión finita y positiva. Dos endomorfismos nilpotentes de V son conjugados si y solamente si tienen el mismo tipo.*

Demostración. Sean f y g dos endomorfismos nilpotentes de V. Si los dos tienen tipo $(n_r, \dots, n_r)$, entonces hay bases ordenadas $\mathscr{B}$ y $\mathscr{B}'$ de V tales que las matrices $[f]_{\mathscr{B}}^{\mathscr{B}}$ y $[g]_{\mathscr{B}'}^{\mathscr{B}'}$ son ambas iguales a

$$\begin{pmatrix} N_{n_1} & & & \\ & N_{n_2} & & \\ & & \ddots & \\ & & & N_{n_r} \end{pmatrix}$$

y esto implica, según la Proposición 6.2.3, que f y g son endomorfismos conjugados.

Recíprocamente, si f y g son conjugados, entonces la Proposición 6.4.4*(i)* nos dice que $\delta_s(f) = \delta_s(g)$ para cada $s \in \mathbb{N}$ y, de acuerdo a la Proposición 6.4.5*(ii)*, de esto se sigue que f y g tienen el mismo tipo. □

6.4.8. Una consecuencia inmediata de los resultados que obtuvimos en esta sección es:

Corolario. *Si V es un espacio vectorial de dimensión finita y positiva n, entonces*

(*i*) *el número de clases de conjugación de elementos nilpotentes de* $\mathsf{End}(V)$ *es finito e igual a p_n, el número de particiones de n, y*

(*ii*) *todos los endomorfismos nilpotentes e indescomponibles de V son conjugados, de manera que son los elementos de exactamente una clase de conjugación de* $\mathsf{End}(V)$.

Demostración. La primera parte es consecuencia de la Proposición 6.4.7, ya que, por un lado, el tipo de un endomorfismo nilpotente de V es, por definición, una partición de n y, por otro, es claro que toda partición de n es el tipo de algún endomorfismo nilpotente de V.

El tipo de un endomorfismo indescomponible de V tiene una única componente —esto se sigue de la Proposición 6.4.3— así que tiene que ser necesariamente igual a (n). Esto implica que, de acuerdo a la primera parte, que todos los endomorfismos nilpotentes de V son conjugados, como afirma la segunda parte de la proposición. □

§5. La forma normal de Jordan

6.5.1. Proposición. *Sea V un espacio vectorial de dimensión finita y positiva n. Si $f : V \to V$ es un endomorfismo indescomponible de V y existe un autovalor λ de f en $\Bbbk$, entonces existe una base ordenada $\mathscr{B}$ de V tal que la matriz de f con respecto a $\mathscr{B}$ es*

$$J_n(\lambda) = \begin{pmatrix} \lambda & & & & \\ 1 & \lambda & & & \\ & 1 & \lambda & & \\ & & \ddots & \ddots & \\ & & & 1 & \lambda \end{pmatrix}$$

La matriz $J_n(\lambda)$ es un ***bloque de Jordan de autovalor λ de tamaño n***. Observemos que cuando $\lambda = 0$ la matriz $J_n(\lambda)$ coincide con la matriz N_n de la sección anterior.

Demostración. Sea $f : V \to V$ un endomorfismo indescomponible de V y sea $\lambda \in \Bbbk$ un autovalor de f. Sea $m_f \in \Bbbk[X]$ el polinomio minimal de f. De acuerdo a la Proposición 6.3.5, hay un polinomio mónico e irreducible $p \in \Bbbk[X]$ y un entero positivo ν tal que $m_f = p^\nu$.

Como λ es una raíz de m_f, claramente tiene que ser también una raíz de p y, entonces, necesariamente debe ser $p = X - \lambda$, ya que p es irreducible. Consideremos el endomorfismo $g = f - \lambda \mathrm{id}_V$ de V. Es $g^\nu = m_f(f) = 0$, así que el polinomio minimal de g divide a X^ν y, en particular, g es nilpotente. Por otro lado, si g fuese descomponible, habría subespacios no nulos y g-invariantes W_1 y W_2 en V tales que $V = W_1 \oplus W_2$: esto es imposible, ya que W_1 y W_2 son también f-invariantes y f es indescomponible por hipótesis.

Vemos así que la Proposición 6.4.3 se aplica a g y que, por lo tanto, existe una base ordenada $\mathscr{B}$ de V tal que $[g]^{\mathscr{B}}_{\mathscr{B}} = N_n$, un bloque de Jordan nilpotente de tamaño n. Se sigue de esto, claro, que

$$[f]^{\mathscr{B}}_{\mathscr{B}} = [g + \lambda \mathrm{id}_V]^{\mathscr{B}}_{\mathscr{B}} = [g]^{\mathscr{B}}_{\mathscr{B}} + \lambda I_n = J_n(\lambda),$$

la matriz descripta en el enunciado. □

6.5.2. Estamos en posición de probar el resultado central de este capítulo, el llamado *Teorema de la forma normal de Jordan*, por Camille Jordan (1838–1922, Francia). Jordan enunció y probó un resultado similar pero distinto, en el contexto de las funciones lineales sobre cuerpos finitos, en su libro *Traité des substitutions et des équations algébriques* [Jor70] de 1870. Es de notar que Jordan no habla nunca de matrices, ni de espacios vectoriales ni de funciones lineales en ese texto: esos conceptos no habían sido inventados aún cuando Jordan escribía.

Teorema. *Sea V un espacio vectorial de dimensión finita y positiva n, sea $f : V \to V$ un endomorfismo de V y supongamos que el polinomio minimal de f se factoriza en $\Bbbk[X]$ como producto de factores lineales. Existen un entero positivo r y un conjunto de r secuencias ordenadas*

$$\{(\lambda_1, n_{1,1}, \dots, n_{1,m_1}), \quad (\lambda_2, n_{2,2}, \dots, n_{2,m_2}), \quad \dots, \quad (\lambda_r, n_{r,r}, \dots, n_{r,m_r})\} \tag{11}$$

unívocamente determinado por f tal que

- *$\lambda_1, \dots, \lambda_r \in \Bbbk$ son escalares distintos dos a dos,*
- *para cada $i \in [\![r]\!]$ es $m_1 \in \mathbb{N}$ y $n_{i,1}, \dots, n_{i,m_1}$ es una secuencia débilmente decreciente de enteros positivos, y*
- *$\sum_{i=1}^{r} \sum_{j=1}^{r_i} n_{i,j} = n$*

para la que existe una base ordenada $\mathscr{B}$ de V tal que la matriz de f con respecto a $\mathscr{B}$ es la matriz diagonal por bloques

$$\begin{pmatrix} J_{n_{1,1}}(\lambda_1) \\ & \ddots \\ && J_{n_{1,m_1}}(\lambda_1) \\ &&& \ddots \\ &&&& \ddots \\ &&&&& \ddots \\ &&&&&& J_{n_{r,1}}(\lambda_r) \\ &&&&&&& \ddots \\ &&&&&&&& J_{n_{r,m_r}}(\lambda_r) \end{pmatrix} \tag{12}$$

Los escalares $\lambda_1, \dots, \lambda_r$ son los autovalores de f. Para cada $i \in [\![n]\!]$ y cada $s \in \mathbb{N}$, la cantidad de bloques de la forma $J_s(\lambda_i)$ que aparecen en la matriz es

$$\mu_s(\lambda_i) = -\delta_{s+1}(f - \lambda_i \mathsf{id}_V) + 2\delta_i(f - \lambda_i \mathsf{id}_V) - \delta_{i-1}(f - \lambda_i \mathsf{id}_V).$$

El número de bloques de autovalor λ_i es igual a la dimensión de $\delta_1(f - \lambda_i \mathsf{id}_V)$.

La matriz (12) que aparece en este teorema es la ***forma normal de Jordan*** del endomorfismo f y el conjunto de r secuencias (11) que la determinan es el ***tipo*** de f.

Demostración. La hipótesis es que existen escalares $\lambda_1, \dots, \lambda_r \in \Bbbk$ distintos dos a dos y enteros positivos $n_1, \dots, n_r$ tales que el polinomio minimal de f es $m_f = (X - \lambda_1)^{n_1} \cdots (X - \lambda_r)^{n_r}$. Sabemos que los autovalores de f son precisamente las raíces del polinomio minimal de f, así que estos r escalares son, sin repeticiones, esos autovalores.

Los r polinomios $q_1 = (X - \lambda_1)^{n_1}, \dots, q_r = (X - \lambda_r)^{n_r}$ son coprimos dos a dos, así que el Teorema de Descomposición Primaria 5.7.1 nos dice que si para cada $i \in [\![r]\!]$ ponemos $V_i = \mathsf{Nu}(f - \lambda_i \mathsf{id}_V)^{n_i}$, entonces $V = V_1 \oplus \cdots \oplus V_r$ y para cada $i \in [\![r]\!]$ el subespacio V_i es f-invariante, no nulo y el polinomio minimal de la restricción $f_{V_i} : V_i \to V_i$ es $m_{f_{V_i}} = (X - \lambda_i)^{n_i}$.

Sea $i \in [\![r]\!]$. Como $(f_{V_i} - \lambda_i \mathsf{id}_V)^{n_i} = m_{f_{V_i}}(f_{V_i}) = 0$, el endomorfismo $f_{V_i} - \lambda_i \mathsf{id}_V : V \to V$ es nilpotente. Si $(n_{i,1}, \dots, n_{i,m_i})$ es su tipo, entonces sabemos que existe una base $\mathscr{B}_i$ de V_i tal que la matriz de $f_{V_i} - \lambda_i \mathsf{id}_V$ con respecto a $\mathscr{B}_i$ es

$$[f_{V_i} - \lambda_i \mathsf{id}_{V_i}]^{\mathscr{B}_i}_{\mathscr{B}_i} = \begin{pmatrix} N_{n_{i,1}} \\ & N_{n_{i,2}} \\ && \ddots \\ &&& N_{n_{i,m_i}} \end{pmatrix},$$

de manera que

$$[f_{V_i}]_{\mathscr{B}_i}^{\mathscr{B}_i} = \begin{pmatrix} N_{n_{i,1}} & & & \\ & N_{n_{i,2}} & & \\ & & \ddots & \\ & & & N_{n_{i,m_i}} \end{pmatrix} + \lambda_i I_{\dim V_i} = \begin{pmatrix} J_{n_{i,1}}(\lambda_i) & & & \\ & J_{n_{i,2}}(\lambda_i) & & \\ & & \ddots & \\ & & & J_{n_{i,m_i}}(\lambda_i) \end{pmatrix}.$$

Si $\mathscr{B} = \mathscr{B}_1 \cup \cdots \cup \mathscr{B}_r$ es la base ordenada de V que se obtiene concatenando las bases $\mathscr{B}_1, \ldots, \mathscr{B}_r$, entonces la matriz $[f]_{\mathscr{B}}^{\mathscr{B}}$ es precisamente la matriz (12) que aparece en el enunciado del teorema. Esto prueba la afirmación de existencia.

Fijemos $j \in [\![r]\!]$ y $s \in \mathbb{N}_0$. Para cada $i \in [\![r]\!]$ el subespacio V_i es $(f - \lambda_j \mathrm{id}_V)$-invariante y la restricción $(f - \lambda_j \mathrm{id}_V)_{V_i}$ claramente coincide con $f_{V_i} - \lambda_j \mathrm{id}_{V_i}$, así que la Proposición 6.4.4*(ii)* nos dice que

$$\delta_s(f - \lambda_j \mathrm{id}_V) = \delta_s(f_{V_1} - \lambda_j \mathrm{id}_{V_1}) + \cdots + \delta_s(f_{V_r} - \lambda_j \mathrm{id}_{V_r}). \tag{13}$$

Ahora bien, si $i \in [\![r]\!]$ es distinto de j, λ_j no es una raíz del polinomio minimal de f_{V_i}, que es $(X - \lambda_i)^{n_i}$, y entonces λ_j no es un autovalor de f_{V_i}: esto implica que el endomorfismo $f_{V_i} - \lambda_j \mathrm{id}_{V_i}$ de V_i es un automorfismo y, en particular, que $\delta_s(f_{V_i} - \lambda_j \mathrm{id}_{V_i}) = 0$. En vista de esto, igualdad (13) nos dice que, de hecho,

$$\delta_s(f - \lambda_j \mathrm{id}_V) = \delta_s(f_{V_j} - \lambda_j \mathrm{id}_{V_1}).$$

Como consecuencia de esto, tenemos que

$$-\delta_{s+1}(f - \lambda_i \mathrm{id}_V) + 2\delta_i(f - \lambda_i \mathrm{id}_V) - \delta_{i-1}(f - \lambda_i \mathrm{id}_V) = -\delta_{s+1}(f_{V_i} - \lambda_i \mathrm{id}_{V_i}) + 2\delta_i(f_{V_i} - \lambda_i \mathrm{id}_{V_i}) - \delta_{i-1}(f_{V_i} - \lambda_i \mathrm{id}_{V_i})$$

y la Proposición 6.4.5*(ii)* nos dice que este último número es el número de veces que el entero s aparece en el tipo de $f_{V_i} - \lambda_i \mathrm{id}_{V_i}$.

□

6.5.3. La prueba que da Jordan de este teorema en [Jor70] es completamente distinta de la que presentamos nosotros. En la tesis [Bre06] de Frédéric Brechenmacher puede encontrarse una discusión detallada de la demostración original. La importancia de este teorema se manifiesta, entre otras formas, en la cantidad de pruebas radicalmente distintas que se han dado de él a lo largo del tiempo: por citar sólo algunas, podemos mencionar las de [Bru87], [Cat62], [Fil71], [FS83], [GW80], [GG96], [Roi99] y [Väl86].

6.5.4. La forma normal de Jordan nos permite resolver el problema que nos habíamos planteado en la sección 6.2 de decidir si dos endomorfismos son conjugados, bajo una hipótesis razonable sobre el cuerpo con el que estamos trabajando:

Proposición. *Supongamos que $\Bbbk$ es un cuerpo algebraicamente cerrado. Dos endomorfismos de un espacio vectorial de dimensión finita y positiva son conjugados si y solamente si tienen el mismo tipo.*

Demostración. HACER. □

6.5.5. Una consecuencia de este resultado es que cualquier invariante bajo de un endomorfismo tiene que poder determinarse a partir de su tipo. Así, por ejemplo, tenemos:

Proposición. *Sea V un espacio vectorial de dimensión finita y positiva n, sea $f : V \to V$ un endomorfismo de V y supongamos que el polinomio minimal de f se factoriza en $\Bbbk[X]$ como producto de factores lineales. Si*

$$(\lambda_1, n_{1,1}, \dots, n_{1,m_1}), \quad (\lambda_2, n_{2,2}, \dots, n_{2,m_2}), \quad \dots, \quad (\lambda_r, n_{r,r}, \dots, n_{r,m_r})$$

es el tipo de f y para cada $i \in [\![r]\!]$ ponemos $n_i = n_{i,1} + \cdots + n_{i,m_i}$ y $\nu_i = \max\{n_{i,1}, \dots, n_{i,m_i}\}$, entonces

- *el polinomio característico de f es $\chi_f = (X - \lambda_1)^{n_1} \cdots (X - \lambda_r)^{n_r}$,*
- *el polinomio minimal de f es $m_f = (X - \lambda_1)^{\nu_1} \cdots (X - \lambda_r)^{\nu_r}$,*
- *el determinante y la traza de f son, respectivamente,*

$$\det(f) = \lambda_1^{n_1} \cdots \lambda_r^{n_r}, \qquad \operatorname{tr}(f) = n_1\lambda_1 + \cdots + n_r\lambda_r,$$

- *los autovalores de f son $\lambda_1, \dots, \lambda_r$ y para cada $i \in [\![r]\!]$ la multiplicidad geométrica de λ_i es m_i y la multiplicidad algebraica es n_i.*

Demostración. En efecto, cada uno de los invariantes listados depende solamente de la clase de conjugación de f y coinciden, por lo tanto, con los de la forma normal de Jordan de f: basta entonces calcularlos para esta, lo que es inmediato. □

Capítulo 7

Espacios con producto interno

§1. Espacios con producto interno

7.1.1. En todo este capítulo escribiremos $\Bbbk$ para referirnos o bien al cuerpo $\mathbb{R}$ de los números reales o bien al cuerpo $\mathbb{C}$ de los números complejos. En cualquiera de los dos casos, si $\lambda \in \Bbbk$, escribimos $\overline{\lambda}$ al número conjugado de λ; si $\Bbbk = \mathbb{R}$, entonces por supuesto es $\overline{\lambda} = \lambda$ para todo $\lambda \in \Bbbk$.

7.1.2. Si V es un espacio vectorial sobre $\Bbbk$, un ***producto interno*** sobre V es una función

$$\langle -, - \rangle : V \times V \to \Bbbk$$

tal que para cada $x, x', y \in V$ y cada $\lambda \in \Bbbk$ se tiene que

(**PI**$_1$) $\langle x + x', y \rangle = \langle x, y \rangle + \langle x', y \rangle$,

(**PI**$_2$) $\langle \lambda x, y \rangle = \lambda \langle x, y \rangle$,

(**PI**$_3$) $\langle x, y \rangle = \overline{\langle y, x \rangle}$,

(**PI**$_4$) $\langle x, x \rangle > 0$ si $x \neq 0$.

Notemos que si $x \in V$, entonces la tercera condición implica que $\langle x, x \rangle = \overline{\langle x, x \rangle}$ y, por lo tanto, el escalar $\langle x, x \rangle$ es un número real cualquiera sea $\Bbbk$: vemos así, en particular, que la última condición de esta definición tiene sentido.

Las condiciones (**PI**$_1$) y (**PI**$_2$) nos dicen que un producto interno es una función lineal de su primera variable, esto es, que para cada $v \in V$ la función

$$u \in V \mapsto \langle u, v \rangle \in \Bbbk$$

es lineal. Si $\Bbbk = \mathbb{R}$, la condición (**PI**$_3$) implica inmediatamente entonces que $\langle -, - \rangle$ es también lineal en su segunda variable y por lo tanto en ese caso se trata, de hecho, de una función bilineal. Si en cambio $\Bbbk = \mathbb{C}$, un producto interno es una función ***semilineal*** de su segundo argumento: esto significa que para cada $u, v, v' \in V$ y cada $\lambda \in \Bbbk$ es $\langle u, v + v' \rangle = \langle u, v \rangle + \langle u, v' \rangle$ y $\langle u, \lambda v \rangle = \overline{\lambda} \langle u, v \rangle$.

En cualquier caso, se sigue de la linealidad en la primera variable que

$$\langle 0,0\rangle = \langle 0\cdot 0,0\rangle = 0\cdot\langle 0,0\rangle = 0$$

y esto, junto con **(PI_4)**, implica que, de hecho, para cada $x\in V$ vale que

$$\langle x,x\rangle = 0 \iff x = 0.$$

7.1.3. Un ***espacio vectorial con producto interno*** es un par ordenado $(V,\langle -,-\rangle)$ en el que V es un espacio vectorial sobre $\Bbbk$ y $\langle -,-\rangle$ es un producto interno sobre V. Salvo en situaciones excepcionales, escribiremos simplemente V en lugar de $(V,\langle -,-\rangle)$ y diremos que V es un espacio con producto interno, dejando implícita la notación para el producto interno.

7.1.4. Ejemplos.

(a) Si $n\in\mathbb{N}$, definimos una función $\langle -,-\rangle : \Bbbk^n\times\Bbbk^n\to\Bbbk$ poniendo, para cada par de vectores $x=(x_1,\dots,x_n)$, $y=(y_1,\dots,y_n)\in\Bbbk^n$,

$$\langle x,y\rangle = x_1\overline{y}_1+\cdots+x_n\overline{y}_n.$$

Es fácil verificar que se trata de un producto interno sobre $\Bbbk^n$, al que llamamos el ***producto interno estándar*** de $\Bbbk^n$. Salvo indicación en contrario, cada vez que consideremos a $\Bbbk^n$ como un espacio vectorial con producto interno será con respecto a a este producto.

Más generalmente, si $\omega=(\omega_1,\dots,\omega_n)$ es un elemento de $\mathbb{R}^n$ con componentes estrictamente positivas, la función $\langle -,-\rangle_\omega : \Bbbk^n\times\Bbbk^n\to\Bbbk$ tal que

$$\langle x,y\rangle_\omega = \omega_1x_1\overline{y}_1+\cdots+\omega_nx_n\overline{y}_n$$

cada vez que $x=(x_1,\dots,x_n)$ e $y=(y_1,\dots,y_n)$ son elementos de $\Bbbk^n$ es un producto interno sobre $\Bbbk^n$, al que llamamos el ***producto interno estándar de peso ω*** sobre $\Bbbk^n$.

(b) Escribamos $\Bbbk^\infty$ al espacio vectorial de las sucesiones $x=(x_i)_{i\geq 1}$ de elementos de $\Bbbk$ cuyas entradas son «casi todas nulas», esto es, aquellas para las que existe $n\geq 1$ tal que $x_i=0$ para todo $i\geq n$. Hay una función $\langle -,-\rangle : \Bbbk^\infty\times\Bbbk^\infty\to\Bbbk$ tal que para cada par de sucesiones $x=(x_i)_{i\geq 1}$ e $y=(y_i)_{i\geq 1}$ de $\Bbbk^\infty$ es

$$\langle x,y\rangle = \sum_{i=1}^{\infty} x_i\overline{y}_i.$$

Notemos que la suma de la derecha tiene sentido, ya que tiene un número finito de términos no nulos. Esta función $\langle -,-\rangle$ es un producto interno sobre $\Bbbk^\infty$.

(c) Sea $\Bbbk[X]$ el espacio vectorial de los polinomios con coeficientes en $\Bbbk$ y sea $I=[a,b]$ un intervalo cerrado y acotado de $\mathbb{R}$ con $a<b$. Hay un producto interno $\langle -,-\rangle : \Bbbk[X]\times\Bbbk[X]\to\Bbbk$ tal que para cada p, $q\in\Bbbk[X]$ se tiene que

$$\langle p,q\rangle = \int_a^b p(t)\overline{q(t)}\,\mathrm{d}t.$$

La verificación de las condiciones (PI$_1$), (PI$_2$) y (PI$_3$) de la definición 7.1.2 es inmediata. Veamos la de (PI$_4$). Sea $p \in \Bbbk[X]$ un polinomio. Es claro que

$$\langle p, p\rangle = \int_a^b |p(t)|^2 \,\mathrm{d}t$$

es un número no negativo, ya que el integrando es real y no negativo. Más aún, si $\langle p, p\rangle = 0$, entonces es nula la integral y, como el integrando es una función continua y no negativa en todo el intervalo I, esto sólo es posible si ese integrando es allí idénticamente nulo: tiene que ser entonces $p(t) = 0$ para cada $t \in I$ y en consecuencia, ya que p es un polinomio y el intervalo I un conjunto infinito, $p = 0$. Esto prueba (PI$_4$)

Más generalmente, si $\omega : I \to \mathbb{R}$ es una función continua, no negativa y no idénticamente nula, es fácil ver que hay un producto interno $\langle -, -\rangle_\omega : \Bbbk[X] \times \Bbbk[X] \to \Bbbk$ sobre $\Bbbk[X]$ tal que para cada $p, q \in \Bbbk[X]$ se tiene que

$$\langle p, q\rangle_\omega = \int_a^b \omega(t)p(t)\overline{q(t)}\,\mathrm{d}t.$$

(d) Sea $I = [a, b]$ un intervalo cerrado y acotado de $\mathbb{R}$ y sea $C(I)$ el espacio vectorial de las funciones $I \to \Bbbk$ que son continuas. Hay una función $\langle -, -\rangle : C(I) \times C(I) \to \Bbbk$ tal que cada vez que f y g son elementos de $C(I)$ se tiene que

$$\langle f, g\rangle = \int_a^b f(t)\overline{g(t)}\,\mathrm{d}x.$$

En efecto, si f y g están en (I) la función $t \in I \mapsto f(t)\overline{g(t)} \in \Bbbk$ es continua y, por lo tanto, está definida su integral sobre I. Es fácil ver que $\langle -, -\rangle$ es un producto interno sobre $C(I)$.

(e) Sea $C(\mathbb{R})$ el espacio vectorial de todas las funciones continuas $\mathbb{R} \to \Bbbk$ y sea $V \subseteq C_\Bbbk(\mathbb{R})$ un subespacio tal que

cualquiera sea $f \in V$ la integral de Riemann impropia $\int_{-\infty}^{\infty} |f(x)|^2 \mathrm{d}x$ es finita.

Por ejemplo, podemos tomar como V al espacio de todas las funciones $f \in C(\mathbb{R})$ para las que existe $M > 0$ tal que $f(x) = 0$ si $|x| > M$.

Mostremos que

si f y g están en V, entonces la integral impropia $\int_{-\infty}^{\infty} f(x)\overline{g(x)}\,\mathrm{d}x$ converge absolutamente. (1)

Sean para ello $f, g \in V$, sean $a = \left(\int_{-\infty}^{\infty} |f(x)|^2 \,\mathrm{d}x\right)^{1/2}$ y $b = \left(\int_{-\infty}^{\infty} |g(x)|^2 \,\mathrm{d}x\right)^{1/2}$, y fijemos fijemos $R > 0$. Si $t \in \mathbb{R}$, entonces

$$\begin{aligned} 0 &\leq \int_{-R}^{R} \big(|f(x)| - t|g(x)|\big)^2 \,\mathrm{d}x \\ &= \int_{-R}^{R} |f(x)|^2 \,\mathrm{d}x - 2t \int_{-R}^{R} |f(x)g(x)| \,\mathrm{d}x + t^2 \int_{-R}^{R} |g(x)| \,\mathrm{d}x. \end{aligned}$$

El último miembro de esta desigualdad es un polinomio en t con coeficientes reales que no toma valores negativos. Su discriminante tiene que ser, por lo tanto, no positivo, y entonces

$$\left(\int_{-R}^{R}|f(x)g(x)|\,\mathrm{d}x\right)^2 \leq \int_{-R}^{R} t^2|f(x)|\,\mathrm{d}x \cdot \int_{-R}^{R} t^2|g(x)|\,\mathrm{d}x \leq a^2b^2.$$

Vemos así que para todo $R > 0$ es $\int_{-R}^{R}|f(x)g(x)|\,\mathrm{d}x \leq ab$ y se sigue de esto que la integral impropia $\int_{-\infty}^{\infty}|f(x)g(x)|\mathrm{d}x$ converge y, por lo tanto, que la integral impropia $\int_{-\infty}^{\infty} f(x)\overline{g(x)}\mathrm{d}x$ converge absolutamente, como queríamos ver.

Como consecuencias de (1) es claro que hay una función $\langle -, - \rangle : V \times V \to \mathbb{R}$ tal que

$$\langle f, g\rangle = \int_{-\infty}^{\infty} f(x)\overline{g(x)}\mathrm{d}x$$

para cada $f, g \in V$. Esta función es un producto interno sobre V. ◇

7.1.5. Proposición. *Sea V un espacio vectorial con producto interno. Si x, $x' \in V$, entonces*

(*i*) $\langle x, y\rangle = 0$ *para todo $x \in V$ si y solamente si $x = 0$, y*

(*ii*) $\langle x, y\rangle = \langle x', y\rangle$ *para todo $y \in V$ si y solamente si $x = x'$.*

Demostración. (*i*) Para ver la necesidad de la condición, notemos que si x la satisface se tiene en particular que $\langle x, x\rangle = 0$, así que $x = 0$. La suficiencia es inmediata.

(*ii*) Esto es consecuencia inmediata de la parte (*i*), ya que $\langle x, y\rangle - \langle x', y\rangle = \langle x - x', y\rangle$ cualquiera sea $y \in V$. □

7.1.6. Podemos restringir el producto interno de un espacio vectorial con producto interno a cualquiera de sus subespacios:

Proposición. *Sea V un espacio vectorial con producto interno y sea W un subespacio de V. La función $\langle -, -\rangle_W : W \times W \to \mathbb{k}$ que se obtiene restringiendo el producto interno $\langle -, -\rangle : V \times V \to \mathbb{k}$ de V a $W \times W$ es un producto interno sobre W.*

Siempre que consideremos a un subespacio de un espacio vectorial con producto interno lo dotaremos con el producto interno construido como en esta proposición.

Demostración. La función $\langle -, -\rangle_W$ satisface cada una de las condiciones de la definición 7.1.2 simplemente porque $\langle -, -\rangle$ lo hace. □

7.1.7. Proposición. *Sean $(V, \langle -, -\rangle_V)$ y $(W, \langle -, -\rangle_W)$ dos espacios con producto interno. Si $V \boxplus W$ es la suma directa externa de V y W, entonces la función*

$$\langle -, -\rangle : (V \boxplus W) \times (V \boxplus W) \to \mathbb{k}$$

tal que para cada de vectores (x, y) y (x', y') de $V \boxplus W$ es

$$\langle (x, y), (x', y')\rangle = \langle x, x'\rangle + \langle y, y'\rangle$$

es un producto interno sobre $V \boxplus W$.

Cuando lo dotamos de este producto interno, llamamos a $V \boxplus W$ la ***suma directa ortogonal*** de los espacios vectoriales con producto interno V y W.

Demostración. HACER. $\square$

§2. Normas y métricas

7.2.1. Si V es un espacio vectorial sobre $\Bbbk$, una ***norma*** sobre V es una función $\|-\| : V \to \mathbb{R}_{\geq 0}$ tal que para cada $x, y \in V$ y cada $\lambda \in \Bbbk$ se satisfacen las condiciones

- ($\mathbf{N}_1$) $\|x\| = 0$ si y solamente si $x = 0$;
- ($\mathbf{N}_2$) $\|\lambda x\| = |\lambda|\|x\|$; y
- ($\mathbf{N}_3$) $\|x + y\| \leq \|x\| + \|y\|$.

Un ***espacio vectorial normado*** es un par ordenado $(V, \|-\|)$ en el que V es un espacio vectorial y $\|-\|$ es una norma sobre V. La desigualdad ($\mathbf{N}_3$) es llamada la ***desigualdad triangular***.

7.2.2. Proposición. *Sea V un espacio vectorial con producto interno. La función*

$$\|-\| : x \in V \mapsto \langle x, x\rangle^{1/2} \in \mathbb{R}.$$

es una norma sobre V y para cada $x, y \in V$ vale la desigualdad

$$|\langle x, y\rangle| \leq \|x\|\|y\|. \tag{2}$$

Más aún, vale aquí la igualdad si y solamente si el conjunto $\{x, y\}$ es linealmente dependiente.

De ahora en adelante consideraremos a todo espacio vectorial con producto interno implícitamente dotado de la norma que nos provee esta proposición, y la llamaremos la ***norma asociada*** a su producto interno. La desigualdad (2) que aparece en esta proposición es la *desigualdad de Cauchy–Bunyakovsky–Schwartz*, por Augustin-Louis Cauchy (1789–1857, Francia), Viktor Bunyakovsky (1804–1889, Rusia) y Hermann Schwarz (1843–1921, Alemania). El primero de estos autores probó la desigualdad en el caso del espacio $\Bbbk^n$ con su producto interno estándar, el segundo para espacios vectoriales de funciones con producto interno dado por una integral, como en el Ejemplo 7.1.4(d), y el tercero probó esencialmente el caso general.

Demostración. Tenemos que verificar las tres condiciones de la definición 7.2.1. Las dos primeras son inmediatas. Para probar la tercera, mostremos antes la desigualdad (2). Sean $x, y \in V$; si $y = 0$, la desigualdad es inmediata, así que podemos suponer que no es ése el caso. Para cada $\lambda \in \Bbbk$ es

$$0 \leq \|x - \lambda y\|^2 = \langle x - \lambda y, x - \lambda y\rangle = \langle x, x\rangle - \lambda\langle y, x\rangle - \overline{\lambda}\big(\langle x, y\rangle - \lambda\langle y, y\rangle\big).$$

Si tomamos $\lambda = \langle x, y\rangle/\langle y, y\rangle$, la expresión entre paréntesis se anula, así que tenemos que

$$0 \leq \langle x, x\rangle - \frac{\langle x, y\rangle\langle y, x\rangle}{\langle y, y\rangle} = \|x\|^2 - \frac{|\langle x, y\rangle|^2}{\|y^2\|},$$

que es equivalente a (2). Más aún, si vale la igualdad, tenemos que $\|x - \lambda y\|^2 = 0$, así que $x = \lambda y$.

Probemos ahora, usando la desigualdad de Cauchy–Bunyakovsky–Schwartz, que la tercera condición de la definición 7.2.1 también se satisface. Si x, $y \in V$, es

$$\begin{aligned}\|x + y\|^2 &= \langle x + y, x + y\rangle = \langle x, x\rangle + \langle x, y\rangle + \langle y, x\rangle + \langle y, y\rangle \\ &= \|x\|^2 + 2\operatorname{Re}\langle x, y\rangle + \|y, y\| \leq \|x\|^2 + 2|\langle x, y\rangle| + \|y, y\| \\ &\leq \|x\|^2 + 2\|x\|\|y\| + \|y, y\| = \big(\|x\| + \|y\|\big)^2\end{aligned}$$

y claramente esto implica que vale la desigualdad triangular. □

7.2.3. Corolario. *Sea V un espacio vectorial con producto interno.*

(i) (Ley del paralelogramo) *Si x, $y \in V$, vale que*

$$\|x + y\|^2 + \|x - y\|^2 = 2\|x\|^2 + 2\|y\|^2.$$

(ii) *Si $\Bbbk = \mathbb{R}$, para cada x, $y \in V$ se tiene que*

$$\langle x, y\rangle = \tfrac{1}{4}\|x + y\|^2 - \tfrac{1}{4}\|x - y\|^2.$$

Si $\Bbbk = \mathbb{C}$, en cambio, es

$$\langle x, y\rangle = \tfrac{1}{4}\|x + y\|^2 - \tfrac{1}{4}\|x - y\|^2 + \tfrac{i}{4}\|x + iy\|^2 - \tfrac{i}{4}\|x - iy\|^2.$$

La ley del paralelogramo tiene una interpretación geométrica muy directa, que puede verse representada gráficamente en la Figura 7.1. Por otro lado, la segunda parte de este corolario nos dice que en un espacio con producto interno el producto interno queda determinado por la norma correspondiente.

Demostración. Para ver la primera parte, sean x e y en V y calculamos:

$$\begin{aligned}\|x + y\|^2 + \|x - y\|^2 &= \langle x + y, x + y\rangle + \langle x - y, x - y\rangle \\ &= \big(\langle x, x\rangle + \langle x, y\rangle + \langle y, x\rangle + \langle y, y\rangle\big) + \big(\langle x, x\rangle - \langle x, y\rangle - \langle y, x\rangle + \langle y, y\rangle\big) \\ &= 2\langle x, x\rangle + 2\langle y, y\rangle \\ &= 2\|x\|^2 + 2\|y\|^2.\end{aligned}$$

Para la segunda parte, suponemos que $\Bbbk = \mathbb{R}$ y vemos que

$$\begin{aligned}\|x + y\|^2 - \|x - y\|^2 &= \langle x + y, x + y\rangle - \|x - y, x - y\| \\ &= \big(\langle x, x\rangle + \langle x, y\rangle + \langle y, x\rangle + \langle y, y\rangle\big) - \big(\langle x, x\rangle - \langle x, y\rangle - \langle y, x\rangle + \langle y, y\rangle\big) \\ &= 4\langle x, y\rangle.\end{aligned}$$

La igualdad que se afirma en la tercera parte de la proposición es consecuencia de un calculo directo similar a este último, que omitimos. □

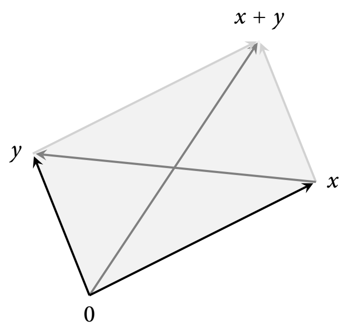


Figura 7.1. La igualdad de la ley del paralelogramo dice que la suma de los cuadrados de las longitudes de las diagonales de un paralelogramo es igual a la suma de los cuadrados de longitudes de los lados.

7.2.4. La primera parte del Corolario 7.2.3 nos dice que la Ley del Paralelogramo es una condición necesaria para que la norma de un espacio normado provenga de un producto interno. El siguiente teorema de Pascual Jordan (1902–1980, Alemania) y John von Neumann (1903–1957, Hungría) publicado en [JvN35] afirma que también es una condición suficiente:

Proposición. *Sea V un espacio vectorial normado. Existe un producto interno sobre V cuya norma asociada es la de V si y solamente si satisface la ley del paralelogramo, esto es, si para todo par de vectores x, y de V se tiene que*

$$\|x+y\|^2 + \|x-y\|^2 = 2\|x\|^2 + 2\|y\|^2.$$

Demostración. Bastará que probemos que la condición es suficiente. Esta condición implica que

$$\|x+x'+y\|^2 = \|(x+y)+x'\|^2 = 2\|x+y\|^2 + 2\|x'\|^2 - \|x+y-x'\|^2$$

y que

$$\|x+x'-y\|^2 = \|x+(x'-y)\|^2 = 2\|x\|^2 + 2\|x'-y\|^2 - \|x-x'+y\|^2.$$

Restando miembro a miembro estas dos igualdades vemos que

$$\|x+x'+y\|^2 - \|x+x'-y\|^2 = 2\|x+y\|^2 + 2\|x'\|^2 - 2\|x\|^2 - 2\|x'-y\|^2$$

e intercambiando aquí los roles de x y de x' concluimos que también

$$\|x+x'+y\|^2 - \|x+x'-y\|^2 = 2\|x'+y\|^2 + 2\|x\|^2 - 2\|x'\|^2 - 2\|x-y\|^2.$$

Sumando ahora miembro a miembro estas dos igualdades llegamos a que

$$\|x+x'+y\|^2 - \|x+x'-y\|^2 = \|x+y\|^2 - \|x-y\|^2 + \|x'+y\|^2 - \|x'-y\|^2 \tag{3}$$

Consideremos primero el caso en que $\Bbbk = \mathbb{R}$. Sea $\langle -, - \rangle : V \times V \to \mathbb{R}$ la función tal que

$$\langle x, y \rangle = \tfrac{1}{4}\|x + y\|^2 - \tfrac{1}{4}\|x - y\|^2$$

para cada x y cada $y \in V$, y verifiquemos que se trata de un producto interno.

- Si $x, x', y \in V$, entonces vale que

$$\langle x + x', y \rangle = \langle x, y \rangle + \langle x', y \rangle,$$

  ya que esta igualdad es, de acuerdo a la definición de la función $\langle -, - \rangle$, equivalente a (3). Esto significa que la condición **(PI$_1$)** de la definición 7.1.2 se satisface.
- Si x e y son elementos de V se tiene que

$$\langle x, y \rangle = \tfrac{1}{4}\|x + y\|^2 - \tfrac{1}{4}\|x - y\|^2 = \tfrac{1}{4}\|y + x\|^2 - \tfrac{1}{4}\|y - x\|^2 = \langle y, x \rangle$$

  y

$$\langle x, x \rangle = \tfrac{1}{4}\|x + x\|^2 - \tfrac{1}{4}\|x - x\|^2 = \|x\|^2 \geq 0, \tag{4}$$

  así que las condiciones **(PI$_3$)** y **(PI$_4$)** también se cumplen.
- Para cada $r \in \mathbb{Q}$ sea $P(r)$ la afirmación

  para todo x, $y \in V$ se tiene que $\langle rx, y \rangle = r\langle x, y \rangle$.

  Queremos mostrar que $P(r)$ vale para todo $r \in \mathbb{Q}$, y lo hacemos en varias etapas:
  - Observemos primero que si r es un número racional cualquiera y vale $P(r)$, entonces para cada $x, y \in V$ se tiene que

$$\langle (r+1)x, y \rangle = \langle rx + x, y \rangle = \langle rx, y \rangle + \langle x, y \rangle = r\langle x, y \rangle + \langle x, y \rangle = (r+1)\langle x, y \rangle,$$

    de manera que también vale $P(r+1)$. Como claramente vale $P(1)$, esto implica que, de hecho, vale $P(r)$ para todo $r \in \mathbb{N}$.
  - Vale $P(r)$ para todo $r \in \mathbb{Q}$ positivo. En efecto, si $r = \frac{p}{q}$ con $p, q \in \mathbb{N}$, para cada $x, y \in V$ se tiene, usando $P(p)$ y $P(q)$, que

$$p\langle x, y \rangle = \langle px, y \rangle = \left\langle q(\tfrac{p}{q}x), y \right\rangle = q\left\langle \tfrac{p}{q}x, y \right\rangle,$$

    así que $\left\langle \frac{p}{q}x, y \right\rangle = \frac{p}{q}\langle x, y \rangle$.
  - Vale $P(0)$, ya que

$$\langle 0x, y \rangle = \tfrac{1}{4}\|0x + y\|^2 - \tfrac{1}{4}\|0x - y\|^2 = \tfrac{1}{4}\|y\|^2 - \tfrac{1}{4}\|y\|^2 = 0 = 0\langle x, y \rangle,$$

    y si vale $P(r)$ para algún $r \in \mathbb{Q}$ entonces también vale $P(-r)$: para todo x e y de V se tiene en ese caso que

$$0 = \langle 0, y \rangle = \langle rx - rx, y \rangle = \langle rx + (-r)x, y \rangle = \langle rx, y \rangle + \langle -rx, y \rangle,$$

    por lo que $\langle -rx, y \rangle = -\langle rx, y \rangle = -r\langle x, y \rangle$. Junto con el paso anterior, esto nos permite concluir que vale $P(r)$ para todo $r \in \mathbb{Q}$, como queríamos.

- Sean otra vez x e y dos elementos de V y mostremos que

$$|\langle x, y\rangle| \leq \|x\|\|y\|. \tag{5}$$

Esto es evidente si $y = 0$, así que podemos suponer que no es ése el caso. En vista de lo que probamos en los dos pasos anteriores, sabemos que cada vez que $r \in \mathbb{Q}$ es

$$0 \leq \|x - ry\|^2 = \langle x - ry, x - ry\rangle = \langle x, x\rangle - 2r\langle x, y\rangle + r^2\langle y, y\rangle.$$

Esto nos dice que el polinomio $\|y\|^2 X^2 - 2\langle x, y\rangle X + \|x\|^2 \in \mathbb{R}[X]$ no toma valores negativos sobre $\mathbb{Q}$: como es una función continua, no toma entonces valores negativos en ningún punto de $\mathbb{R}$ y, por lo tanto, su discriminante es no negativo, esto es, $\langle x, y\rangle^2 - \|x\|^2\|y\|^2 \leq 0$. La desigualdad (5) es consecuencia inmediata de esto.

- Si $x, y \in V$, entonces la función $\zeta : t \in \mathbb{R} \mapsto \langle tx, y\rangle \in \mathbb{R}$ es continua. Para verlo basta observar que, de acuerdo a lo que sabemos ya de la función $\langle -, -\rangle$, se tiene que

$$|\zeta(s) - \zeta(t)| = \left|\langle sx, y\rangle - \langle tx, y\rangle\right| = \left|\langle (s-t)x, y\rangle\right| \leq \|(s-t)x\|\|y\| \leq |s-t|\|x\|\|y\|.$$

- Sean x e y dos elementos de V. Sabemos, gracias a los dos pasos anteriores, que la función $t \in \mathbb{R} \mapsto \langle tx, y\rangle - t\langle x, y\rangle \in \mathbb{R}$ es continua y que se anula sobre $\mathbb{Q}$. Esto implica, por supuesto, que es idénticamente nula y, por lo tanto, que para todo $t \in \mathbb{R}$ vale que $\langle tx, y\rangle = t\langle x, y\rangle$. La función $\langle -, -\rangle$ satisface entonces la condición **(PI$_2$)**.

Concluimos de esta forma que la función $\langle -, -\rangle$ es un producto interno sobre V. De acuerdo a la igualdad (4), la norma asociada a este producto interno coincide con la norma de V: esto completa la prueba de la proposición en el caso que estamos considerando en el que $\Bbbk = \mathbb{R}$.

HACER: El caso complejo. □

7.2.5. Ejemplo. La función $\|-\| : (x, y) \in \Bbbk^2 \mapsto |x| + |y| \in \mathbb{R}_{\geq 0}$ es una norma sobre $\Bbbk^2$. Si $e_1 = (1, 0)$ y $e_2 = (0, 1)$ son los dos vectores de la base estándar de $\Bbbk^2$, entonces

$$\|e_1 + e_2\|^2 + \|e_1 - e_2\|^2 = 8 \neq 4 = \|e_1\|^2 + 2\|e_2\|^2.$$

Vemos así que esta norma $\|-\|$ no satisface la ley de paralelogramo y, en consecuencia, que no existe ningún producto interno sobre V que la tenga como norma asociada. ◇

7.2.6. Si V es un espacio vectorial, una función $d : V \times V \to \mathbb{R}_{\geq 0}$ es una ***métrica*** sobre V si para cada $x, y, z \in V$ se tiene que

(M$_1$) $d(x, y) = d(y, x)$;

(M$_2$) $d(x, y) = 0$ si y solamente si $x = y$; y

(M$_3$) $d(x, z) \leq d(x, y) + d(y, z)$.

Decimos que esa métrica es ***invariante*** si

(M$_4$) $d(x, y) = d(x + z, y + z)$

y que es ***homogénea*** si para cada $\lambda \in \Bbbk$ es

(**M**$_5$) $d(\lambda x, \lambda y) = |\lambda|\, d(x, y)$.

7.2.7. Proposición. *Sea V un espacio vectorial con producto interno y sea $\|-\| : V \to \mathbb{R}$ la norma asociada. La función*

$$d : (v, w) \in V \times V \mapsto \|v - w\| \in \mathbb{R}_{\geq 0}$$

es una métrica invariante y homogénea sobre V.

Demostración. Esto es una consecuencia inmediata de la definición de la función d y de las propiedades de la norma $\|-\|$. □

§3. Ortogonalidad

7.3.1. Sea V un espacio con producto interno. Si x e y son dos vectores de V, decimos que x e y son ***ortogonales*** si $\langle u, v\rangle = 0$, y en ese caso escribimos $u \perp v$. Es inmediato que esto define una relación $\perp$ entre los elementos de V que es simétrica. Además, tenemos que

Lema. *Sea V un espacio vectorial con producto interno.*

(*i*) *Todo vector de V es ortogonal con* 0.

(*ii*) *Si $x \in V$ es tal que $x \perp y$ para todo $y \in V$, entonces $x = 0$.*

Demostración. La primera afirmación es inmediata, y la segunda es consecuencia de la primera parte de la Proposición 7.1.5. □

7.3.2. Intuitivamente, dos vectores son ortogonales si el ángulo que forman es recto. Esto se ve reflejado en que vale el teorema de Pitágoras:

Proposición. *Sea V un espacio con producto interno. Si x, $y \in V$, entonces*

(*i*) *si $x \perp y$, entonces* $\|x + y\|^2 = \|x\|^2 + \|y\|^2$.

Más generalmente, vale que

(*ii*) $\|x + y\|^2 = \|x\|^2 + \|y\|^2 + 2\,\mathrm{Re}\langle x, y\rangle$.

Demostración. La primera afirmación es consecuencia inmediata de la segunda, así que basta probar esta última, que a su vez sigue de un cálculo directo: si x, $y \in V$, la definición de la norma implica que

$$\begin{aligned}
\|x + y\|^2 &= \langle x + y, x + y\rangle \\
&= \langle x, x\rangle + \langle x, y\rangle + \langle y, x\rangle + \langle y, y\rangle \\
&= \|x\|^2 + \|y\|^2 + \langle x, y\rangle + \overline{\langle x, y\rangle} \\
&= \|x\|^2 + \|y\|^2 + 2\,\mathrm{Re}\langle x, y\rangle.
\end{aligned}$$

□

7.3.3. Motivados por la Proposición 7.3.2, hacemos la siguiente definición: si V es un espacio vectorial con producto interno y x e y son dos vectores de V, el ***ángulo*** entre x e y es el único número real $\theta(x,y) \in [0,\pi]$ tal que

$$\cos\theta(x,y) = \frac{\operatorname{Re}\langle x,y\rangle}{\|x\|\|y\|}. \tag{6}$$

Esta definición tiene sentido: de la desigualdad de Cauchy-Bunyakovsky-Schwartz sabemos que

$$|\operatorname{Re}\langle x,y\rangle| \leq |\langle x,y\rangle| \leq \|x\|\|y\|,$$

así que el cociente que aparece a la derecha en la definición (6) es un elemento del intervalo $[-1,1]$ y, por lo tanto, es el coseno de exactamente un número θ en el intervalo $[0,\pi]$. Es inmediato que dos vectores son ortogonales si y solamente si el ángulo entre ellos es $\pi/2$.

Usando esta definición, podemos reinterpretar la segunda parte de la Proposición 7.3.2 como el *teorema del coseno* de la geometría euclídea:

Proposición. *Sea V un espacio vectorial con producto interno. Si x e y son dos vectores de V y $\theta(x,y)$ es el ángulo entre x e y, entonces*

$$\|x-y\|^2 = \|x\|^2 + \|y^2\| - 2\cos\theta(x,y).$$

Demostración. Basta reemplazar y por $-y$ en la igualdad de la Proposición 7.3.2*(ii)* y usar la definición de $\theta(x,y)$. □

7.3.4. Si V es un espacio con producto interno y A es un subconjunto de V, entonces decimos que

- A es ***ortogonal*** si para cada par $x, y \in A$ de elementos distintos se tiene que $x \perp y$, y que
- A es ***ortonormal*** si es ortogonal y además $\|x\| = 1$ para cada $x \in A$.

7.3.5. Hay una relación muy útil entre la ortogonalidad y la independencia lineal:

Proposición. *Sea V un espacio con producto interno y sea A un subconjunto de V.*

(*i*) *Si A es ortogonal y $0 \notin A$, entonces A es linealmente independiente.*

(*ii*) *Si A es ortonormal y si $x = \lambda_n x_n + \cdots + \lambda_n x_n$, con $x_1, \ldots, x_n \in A$ distintos dos a dos y $\lambda_1, \ldots, \lambda_n \in \Bbbk$, entonces $\lambda_j = \langle x, x_j\rangle$ para todo $j \in [\![n]\!]$.*

(*iii*) *Si A es ortonormal, entonces para cada $x \in \langle A\rangle$ el conjunto $A_x = \{y \in A : \langle x,y\rangle \neq 0\}$ es finito, y vale que $x = \sum_{y\in A_x}\langle x,y\rangle y$.*

Demostración. (*i*) Sean $n \geq 1$, sean $x_1, \ldots, x_n \in A$ distintos dos a dos y sean $\lambda_1, \ldots, \lambda_n \in \Bbbk$ tales que $\sum_{i=1}^n \lambda_i x_i = 0$. Si $j \in [\![n]\!]$, entonces

$$0 = \langle 0, x_j\rangle = \left\langle \sum_{i=1}^n \lambda_i x_i, x_j \right\rangle = \sum_{i=1}^n \lambda_i \langle x_i, x_j\rangle.$$

Como A es ortogonal, en esta última suma el único término que puede ser no nulo es aquél en el que i es igual a j, así que tenemos que $\lambda_j\langle a_j, a_j\rangle = 0$. Como $0 \notin A$, es $\langle a_j, a_j\rangle \neq 0$ y entonces $\lambda_j = 0$. Esto vale para cada $j \in [\![n]\!]$, así que podemos concluir que A es linealmente independiente.

(*ii*) Si $j \in [\![n]\!]$, entonces

$$\langle x, x_j \rangle = \left\langle \sum_{i=1}^{n} \lambda_i x_i, x_j \right\rangle = \sum_{i=1}^{n} \lambda_i \langle x_i, x_j \rangle.$$

Como A es ortonormal, el único término de esta suma que es posiblemente no nulo es el que tiene i igual a j, que es igual a $\lambda_j \langle x_j, x_j \rangle = \lambda_j$. Esto prueba lo que queremos.

(*iii*) Sea $x \in \langle A \rangle$, de manera que existen $n \geq 0$, $x_1, \ldots, x_n \in A$ y $\lambda_1, \ldots, \lambda_n \in \Bbbk$ tales que $x = \sum_{i=1}^{n} \lambda_i x_i$, y supongamos, sin pérdida de generalidad, que $\lambda_i \neq 0$ para cada $i \in [\![n]\!]$.

Si $y \in A$, entonces $\langle x, y \rangle = \sum_{i=1}^{n} \lambda_i \langle x_i, y \rangle$. Como A es ortonormal, esta suma tiene todos sus términos nulos si $y \notin \{x_1, \ldots, x_n\}$. Así, vemos que A_x es un subconjunto de $\{x_1, \ldots, x_n\}$ y, en particular, que es finito. Para terminar, es suficiente con mostrar que $\lambda_i = \langle x, x_i \rangle$ para cada $i \in [\![n]\!]$: esto es consecuencia inmediata de la afirmación (*ii*). □

7.3.6. Corolario. *Sea V un espacio vectorial con producto interno y de dimensión finita n, y supongamos que $\mathscr{B} = \{x_1, \ldots, x_n\}$ es una base ortonormal de V. Para cada $x \in V$ es*

$$x = \sum_{i=1}^{n} \langle x, x_i \rangle x_i.$$

Demostración. Esto es un caso particular de la tercera parte de la Proposición 7.3.5*(iii)*. □

7.3.7. Es fácil construir conjuntos ortonormales:

Proposición. *Sea V un espacio vectorial con producto interno. Si $n \geq 1$ y $x_1, \ldots, x_n$ son n vectores de V tal que el conjunto $\{x_1, \ldots, x_n\}$ tiene n elementos y es linealmente independiente, entonces hay vectores $y_1, \ldots, y_n$ en V tales que*

- *el conjunto $\{y_1, \ldots, y_n\}$ es ortonormal,*
- *para todo $i \in [\![n]\!]$ se tiene que $\langle x_1, \ldots, x_i \rangle = \langle y_1, \ldots, y_i \rangle$, y*
- *para todo $i \in [\![n]\!]$ es $\langle x_i, y_i \rangle > 0$.*

Estas condiciones determinan unívocamente a los vectores $y_1, \ldots, y_n$.

Demostración. Probemos la existencia haciendo inducción con respecto a n.

Supongamos primero que $n = 1$. Como $\{x_1\}$ es linealmente independiente, es $x_1 \neq 0$ y entonces $\|x_1\| \neq 0$: podemos considerar entonces el vector $y_1 = x_1/\|x_1\|$. Es inmediato que $\|y_1\| = 1$, de manera que $\{y_1\}$ es un conjunto ortonormal, que $\langle x_1 \rangle = \langle y_1 \rangle$ y que $\langle x_1, y_1 \rangle > 0$. La afirmación de la proposición es por lo tanto cierta en este caso.

Supongamos ahora que $n > 1$. Como $\{x_1, \ldots, x_{n-1}\}$ es linealmente independiente y tiene $n - 1$ elementos, podemos suponer inductivamente que hay un conjunto ortonormal $\{y_1, \ldots, y_{n-1}\}$ tal que para cada $i \in [\![n-1]\!]$ es $\langle x_1, \ldots, x_i \rangle = \langle y_1, \ldots, y_i \rangle$. Sea

$$\tilde{y}_n = x_n - \sum_{i=1}^{n-1} \langle x_n, y_i \rangle y_i.$$

Si fuese $\tilde{y}_n = 0$, tendríamos que

$$x_n = \sum_{i=1}^{n-1} \langle x_n, y_i \rangle y_i \in \langle y_1, \dots, y_{n-1} \rangle = \langle x_1, \dots, x_{n-1} \rangle,$$

lo que es imposible, ya que $\{v_1, \dots, v_n\}$ es linealmente independiente. Podemos entonces considerar el vector $y_n = \tilde{y}_n / \|y_n\|$, que tiene norma unitaria. Si $j \in [\![n-1]\!]$, entonces

$$\begin{aligned} \|\tilde{y}_n\| \langle y_n, y_j \rangle &= \left\langle x_n - \sum_{i=1}^{n-1} \langle x_n, y_i \rangle y_i, y_j \right\rangle = \langle x_n, y_j \rangle - \sum_{i=1}^{n-1} \langle x_n, y_i \rangle \langle y_i, y_j \rangle \\ &= \langle x_n, y_j \rangle - \langle x_n, y_j \rangle = 0, \end{aligned}$$

y esto, junto con el hecho de que $\{y_1, \dots, y_{n-1}\}$ es un conjunto ortonormal, implica que el conjunto $\{y_1, \dots, y_{n-1}, y_n\}$ es ortonormal. Como es claro que

$$\langle x_1, \dots, x_n \rangle = \langle y_1, \dots, y_n \rangle,$$

esto completa la inducción.

HACER: Unicidad. □

7.3.8. La demostración que dimos de la Proposición 7.3.7 nos da un procedimiento efectivo para obtener conjuntos ortonormales, llamado ***procedimiento de ortogonalización de Gram–Schmidt***, por Jørgen Pedersen Gram (1850–1916, Dinamarca) y Erhard Schmidt (1876–1959, Alemania). Empezando con n vectores $x_1, \dots, x_n$ linealmente independientes calculamos, en orden,

$$\begin{aligned} &\tilde{y}_1 = x_1, && y_1 = \tilde{y}_1 / \|\tilde{y}_1\|, \\ &\tilde{y}_2 = x_2 - \langle x_2, y_1 \rangle y_1, && y_2 = \tilde{y}_2 / \|\tilde{y}_2\|, \\ &\tilde{y}_3 = x_3 - \langle x_3, y_1 \rangle y_1 - \langle x_3, y_2 \rangle y_2, && y_3 = \tilde{y}_3 / \|\tilde{y}_3\|, \\ &\dots, && \dots, \\ &\tilde{y}_n = x_n - \langle x_n, y_1 \rangle y_1 - \langle x_n, y_2 \rangle y_2 - \dots - \langle x_n, y_{n-1} \rangle y_{n-1}, && y_n = \tilde{y}_n / \|\tilde{y}_n\|. \end{aligned}$$

Al terminar, el conjunto $\{y_1, \dots, y_n\}$ es ortonormal y genera el mismo espacio que $\{x_1, \dots, x_n\}$.

7.3.9. Ejemplos.

(a) Consideremos al espacio vectorial $\Bbbk^3$ dotado de su producto interno estándar y los tres vectores

$$x_1 = (1, 1, 0), \quad x_2 = (2, 2, 2), \quad x_3 = (0, 1, 1).$$

Siguiendo el procedimiento de ortogonalización descripto arriba, encontramos:

$$
\begin{aligned}
y_1 &= x_1/\|x_1\| = (\tfrac{1}{\sqrt{2}}, \tfrac{1}{\sqrt{2}}, 0),\\
\tilde{y}_2 &= x_2 - \langle x_2, y_1\rangle y_1 = (0,0,2),\\
y_2 &= \tilde{y}_2/\|y_2\| = (0,0,1),\\
\tilde{y}_3 &= x_3 - \langle x_3, y_1\rangle y_1 - \langle x_3, y_2\rangle y_2 = (-\tfrac{1}{2}, \tfrac{1}{2}, 0),\\
y_3 &= \tilde{y}_3/\|y_3\| = (-\tfrac{1}{\sqrt{2}}, \tfrac{1}{\sqrt{2}}, 0).
\end{aligned}
$$

El conjunto $\{(\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}, 0), (0,0,1), (-\frac{1}{\sqrt{2}}, \frac{1}{\sqrt{2}}, 0)\}$ es entonces ortonormal y genera el mismo subespacio que $\{x_1, x_2, x_3\}$.

(b) Sea ahora $\mathbb{R}[X]$ el espacio vectorial real de los polinomios con coeficientes en $\mathbb{R}$ y consideremos sobre él el producto interno $\langle -, -\rangle : \mathbb{R}[X] \times \mathbb{R}[X] \to \mathbb{R}$ tal que

$$\langle p, q\rangle = \frac{2}{\pi}\int_{-1}^{1} f(t)g(t)\sqrt{1-t^2}\,\mathrm{d}t$$

para cada $p, q \in \mathbb{R}[X]$; observemos que esto es un caso particular de la construcción hecha en el Ejemplo 7.1.4(c) con $I = [-1,1]$ y $\omega : t \in I \mapsto \frac{2}{\pi}\sqrt{1-t^2} \in \mathbb{R}$. Un cálculo directo muestra que si realizamos el procedimiento de ortogonalización de Gram–Schmidt empezando con los polinomios

$$1, \quad X, \quad X^2, \quad X^3, \quad X^4, \quad \dots$$

obtenemos los polinomios

$$1, \quad 2X, \quad 4X^2-1, \quad 8X^3-4X, \quad 16X^4-12X^2+1, \quad 32X^5-32X^3+6X, \quad \dots$$

Estos polinomios son precisamente los polinomios de Chebyshev que encontramos en el Ejemplo 5.3.10(b) del Capítulo 5. Probemos esto.

Para cada $n \in \mathbb{N}_0$ sea U_n el n-ésimo polinomios de Chebyshev, de manera que

$$U_0 = 1, \qquad U_1 = 2X \tag{7}$$

y

$$U_n = 2XU_{n-1} - U_{n-2} \text{ si } n \geq 2. \tag{8}$$

Si $n, m \in \mathbb{N}_0$, entonces

$$\langle U_n, U_m\rangle = \frac{2}{\pi}\int_{-1}^{1} U_n(t)U_m(t)\sqrt{1-t^2}\,\mathrm{d}t. \tag{9}$$

Como vimos en el Ejemplo 5.3.10(b), vale que si $\theta \in (0,\pi)$ es

$$U_n(\cos\theta) = \frac{\sin(n+1)\theta}{\sin\theta}.$$

y entonces, poniendo $t = \cos\theta$ en la integral (9), vemos que

$$\begin{aligned}\langle U_n, U_m\rangle &= \frac{2}{\pi}\int_0^\pi U_n(\cos\theta)U_m(\cos\theta)\sin^2\theta\,\mathrm{d}\theta \\ &= \frac{2}{\pi}\int_0^\pi \sin(n+1)\theta\cdot\sin(m+1)\theta\,\mathrm{d}\theta \\ &= \frac{1}{\pi}\int_0^\pi \cos(n-m)\theta\,\mathrm{d}\theta - \frac{1}{\pi}\int_0^\pi \cos(n+m+2)\theta\,\mathrm{d}\theta \end{aligned} \tag{10}$$

Ahora bien, es fácil ver que para cada $k \in \mathbb{Z}$ se tiene que

$$\int_0^\pi \cos k\theta\,\mathrm{d}\theta = \begin{cases} \pi, & \text{si } k = 0; \\ 0, & \text{en caso contrario;} \end{cases}$$

y usando esto en (10) concluimos inmediatamente que

$$\langle U_n, U_m\rangle = \begin{cases} 1, & \text{si } n = m; \\ 0, & \text{en caso contrario.} \end{cases}$$

Esto prueba que el conjunto $\{U_n : n \in \mathbb{N}_0\}$ es ortonormal.

Por otro lado, una demostración inductiva a partir de (7) y (8) muestra inmediatamente que para todo $n \in \mathbb{N}_0$ el grado de U_n es n y que el coeficiente de X^n en U_n es 2^n. **Hacer:** Terminar.

◇

7.3.10. La consecuencia más importante de la Proposición 7.3.7 es la siguiente:

Corolario. *Un espacio vectorial con producto interno y de dimensión finita tiene una base ortonormal.*

Demostración. Si V es un espacio vectorial con producto interno y de dimensión finita y $\mathscr{B}$ es una base de V, la proposición nos da un conjunto ortonormal $\mathscr{B}'$ con la misma cantidad de elementos que $\mathscr{B}$ y que, entre otras cosas, tiene la propiedad de que $\langle\mathscr{B}'\rangle = \langle\mathscr{B}\rangle = V$. Esto implica que $\mathscr{B}'$ es una base de V. Como $\mathscr{B}'$ es ortonormal, esto prueba el corolario. □

7.3.11. Vimos como un producto interno nos permite «medir» la distancia entre dos elementos de un espacio vectorial. Extendemos ahora esa idea para poder hablar de la distancia de un punto a un conjunto. Si V es un espacio vectorial con producto interno, S un subconjunto no vacío de V y $x \in V$, la ***distancia de x a S*** es el número

$$d(x, S) = \inf\{d(x, y) : y \in S\}.$$

Observemos que esta definición tiene sentido, ya que el subconjunto $\{d(x, y) : y \in S\}$ de $\mathbb{R}$ no es vacío y está acotado inferiormente por 0, así que posee un ínfimo.

7.3.12. Proposición. *Sea V un espacio vectorial con producto interno, sea S un subespacio de dimensión finita de V y sea $x \in V$. Si $\mathscr{B} = \{y_1, \dots, y_n\}$ es una base ortonormal de S y*

$$x_S = \sum_{j=1}^{n}\langle x, y_j\rangle y_j,$$

entonces

(*i*) $x - x_S \perp S$,

(*ii*) $d(x, S) = d(x, x_S)$, *y*

(*iii*) $d(x, x_S) < d(x, y)$ *para todo* $y \in S \setminus \{x_S\}$.

Esta proposición nos dice que el vector x_S es un punto de S tal que $d(x, S) = d(x, x_S)$ y que es, más aún, el único con esa propiedad.

Demostración. (*i*) Sea $y \in S$. Como $\mathscr{B}$ es una base ortonormal de S, es $y = \sum_{i=1}^{n} \langle y, y_i \rangle y_i$ y entonces

$$\begin{aligned}
\langle x - x_S, y \rangle &= \Big\langle x - \sum_{j=1}^{n} \langle x, y_j \rangle y_j, \sum_{i=1}^{n} \langle y, y_i \rangle y_i \Big\rangle \\
&= \sum_{i=1}^{n} \overline{\langle y, y_i \rangle} \langle x, y_i \rangle - \sum_{j=1}^{n} \sum_{i=1}^{n} \langle x, y_j \rangle \overline{\langle y, y_i \rangle} \langle y_j, y_i \rangle \\
&= \sum_{i=1}^{n} \overline{\langle y, y_i \rangle} \langle x, y_i \rangle - \sum_{j=1}^{n} \langle x, y_j \rangle \overline{\langle y, y_j \rangle} = 0.
\end{aligned}$$

Vemos así que $x - x_S \perp S$.

(*ii*) Como $x_S \in S$, es claro que $d(x, S) \leq d(x, x_S)$. Si, por otro lado, es $y \in S$, tenemos que

$$\begin{aligned}
d(x, y)^2 &= \|x - y\|^2 \\
&= \|(x - x_S) + (x_S - y)\|^2
\end{aligned}$$

y, como $x - x_S \perp x_S - y$ porque $x_S - y \in S$, esto es

$$\begin{aligned}
&= \|x - x_S\|^2 + \|x_S - y\|^2 \\
&\geq \|x - x_S\|^2,
\end{aligned}$$

de manera que $d(x, y) \geq d(x, x_S)$. Vemos así que $d(x, S) \geq d(x, x_S)$.

(*iii*) Si $y \in S \setminus \{x_S\}$, entonces $d(x_S, y) > 0$ y, como recién,

$$d(x, y)^2 = d(x, x_S)^2 + d(x_S, y)^2 > d(x, x_S)^2,$$

así que $d(x, y) > d(x, x_S)$. □

§4. Complementos ortogonales

7.4.1. Si V es un espacio vectorial con producto interno y S es un subconjunto arbitrario de V, el ***complemento ortogonal*** de S es el subconjunto

$$S^\perp = \{x \in V : x \perp s \text{ para todo } s \in S\}.$$

7.4.2. Proposición. *Sea V un espacio vectorial con producto interno.*

(*i*) *Si S es un subconjunto de V, entonces $S^\perp$ es un subespacio de V.*

(*ii*) *Se tiene que $0^\perp = V$ y $V^\perp = 0$.*

(*iii*) *Si S y T son subconjuntos de V y $S \subseteq T$, entonces $T^\perp \subseteq S^\perp$.*

(*iv*) *Si S es un subconjunto de V, entonces $\langle S\rangle^\perp = S^\perp$.*

(*v*) *Si S y T son subespacios de V, entonces $(S+T)^\perp = S^\perp \cap T^\perp$.*

Demostración. (*i*) Sean $x, y \in S^\perp$ y $a, b \in \Bbbk$. Para cada $s \in S$, es

$$\langle ax + by, s\rangle = a\langle x, s\rangle + b\langle y, s\rangle = 0,$$

porque ambos términos se anulan. Esto nos dice que $ax + by \in S^\perp$. Como $S^\perp$ no es vacío, ya que contiene a 0, se trata de un subespacio de V.

(*ii*) Todo vector es ortogonal a 0, así que $V \subseteq 0^\perp$. La primera igualdad es entonces inmediata. Por otro lado, si $x \in V^\perp$, entonces $\langle x, y\rangle = 0$ para todo $y \in V$ y la Proposición 7.1.5 nos dice que $x = 0$. Como por supuesto $0 \in V^\perp$, esto prueba la segunda igualdad del enunciado.

(*iii*) Supongamos que S y T son subconjuntos de V y que $S \subseteq T$. Si $y \in T^\perp$, entonces para cada $s \in S$ se tiene que $y \perp s$, ya que $s \in T$ y, en consecuencia, $x \in S^\perp$. Así, es $T^\perp \subseteq S^\perp$, como queremos.

(*iv*) Como $S \subseteq \langle S\rangle$, la parte (*iii*) implica que $\langle S\rangle^\perp \subseteq S^\perp$. Sea, por otro lado, $x \in S^\perp$ y sea $y \in \langle S\rangle$. Existen entonces $n \geq 0$, $s_1, \dots, s_n \in S$ y $\lambda_1, \dots, \lambda_n \in \Bbbk$ tales que $y = \sum_{i=1}^n \lambda_i s_i$ y, en consecuencia,

$$\langle x, y\rangle = \Big\langle x, \sum_{i=1}^n \lambda_i s_i\Big\rangle = \sum_{i=1}^n \overline{\lambda}_i \langle x, s_i\rangle = 0$$

porque $x \in S^\perp$. Esto muestra que $S^\perp \subseteq \langle S\rangle^\perp$.

(*v*) Sean ahora S y T dos subespacios de V. Como S y T están contenidos en $S+T$, la parte (*ii*) implica que $(S+T)^\perp \subseteq S^\perp$ y $(S+T)^\perp \subseteq T^\perp$, así que $(S+T)^\perp \subseteq S^\perp \cap T^\perp$. Por otro lado, si $x \in S^\perp \cap T^\perp$ e $y \in S+T$, de manera que existen $s \in S$ y $t \in T$ tales que $y = s + t$, entonces

$$\langle x, y\rangle = \langle x, s + t\rangle = \langle x, s\rangle + \langle x, t\rangle = 0.$$

Esto nos dice que $x \in (S+T)^\perp$ y, en definitiva, que $S^\perp \cap T^\perp \subseteq (S+T)^\perp$. □

7.4.3. Si S es un subconjunto de un espacio vectorial con producto interno, escribimos $S^{\perp\perp} = (S^\perp)^\perp$ y $S^{\perp\perp\perp} = (S^{\perp\perp})^\perp$. Observemos que $S^{\perp\perp\perp}$ coincide con $(S^\perp)^{\perp\perp}$.

Proposición. *Sea V un espacio vectorial con producto interno. Si S es un subconjunto de V, entonces*

(*i*) $S \cap S^\perp \subseteq 0$,

(*ii*) $S \subseteq S^{\perp\perp}$,

(*iii*) $S^\perp = S^{\perp\perp\perp}$.

Demostración. (*i*) Si $x \in S \cap S^\perp$, entonces $\langle x, x\rangle = 0$, ya que $x \in S$ y $x \in S^\perp$, y por lo tanto $x = 0$.

(*ii*) Si $s \in S$, para cada $t \in S^\perp$ es $s \perp t$. Esto dice que $s \in (S^\perp)^\perp = S^{\perp\perp}$.

(*iii*) Acabamos de probar que $S \subseteq S^{\perp\perp}$, así que usando la Proposición 7.4.2*(iii)* vemos que $S^{\perp\perp\perp} = (S^{\perp\perp})^\perp \subseteq S^\perp$. Por otro lado, según la parte (*ii*), $S^\perp \subseteq (S^\perp)^{\perp\perp} = S^{\perp\perp\perp}$. □

7.4.4. En la situación de la Proposición 7.4.3*(ii)*, la inclusión es en general estricta. Vale la igualdad, sin embargo, en un caso importante:

Proposición. *Sea V un espacio vectorial con producto interno. Si S es un subespacio de V de dimensión finita, entonces $V = S \oplus S^{\perp}$ y $S = S^{\perp\perp}$.*

Demostración. Sea $x \in V$. De acuerdo a la Proposición 7.3.12, existe $x_S \in S$ tal que $x - x_S \perp S$, esto es, tal que $x - x_S \in S^{\perp}$. Entonces $x = x_S + (x - x_S) \in S + S^{\perp}$ y vemos que $V = S + S^{\perp}$. Como sabemos que $S \cap S^{\perp} = 0$, esto implica que $V = S \oplus S^{\perp}$.

Para probar que $S = S^{\perp\perp}$ basta mostrar, en vista de la Proposición 7.4.3*(ii)*, que $S^{\perp\perp} \subseteq S$. Sea entonces $x \in S^{\perp\perp}$. Como $V = S \oplus S^{\perp}$, existen $s \in S$ y $t \in S^{\perp}$ tales que $x = s + t$. Como $x \in S^{\perp\perp}$, es

$$0 = \langle x, t\rangle = \langle s + t, t\rangle = \langle s, t\rangle + \langle t, t\rangle = \langle t, t\rangle,$$

así que $t = 0$ y, por lo tanto, $x = s + t = s \in S$. □

7.4.5. Corolario. *Sea V un espacio vectorial con producto interno y de dimensión finita. Si $S \subseteq V$ es un subespacio, entonces*

$$\dim V = \dim S + \dim S^{\perp}.$$

Demostración. Esto es consecuencia de que, de acuerdo a la proposición, $V = S \oplus S^{\perp}$. □

7.4.6. Ejemplo. Demos un ejemplo de una situación en la que la inclusión de la Proposición 7.4.3*(ii)* es propia. Sea ℓ_2 el conjunto de todas las sucesiones $(x_i)_{i\geq 1}$ de elementos de $\Bbbk$ tales que la serie $\sum_{i=1}^{\infty}|x_i|^2$ converge a un número finito. Esto es un subconjunto del espacio vectorial complejo de todas las sucesiones $(x_i)_{i\geq 1}$ y, de hecho, se trata de un subespacio: es inmediato que si $\xi \in \ell_2$ y $\lambda \in \mathbb{C}$ se tiene que $\lambda\xi \in \ell_2$, así que bastará que mostremos que ℓ_2 es cerrado para la suma.

Sean $\xi = (x_i)_{i\geq 1}$ y $\zeta = (y_i)_{i\geq 1}$ dos elementos de ℓ_2 y consideremos los números $a = \sum_{i=1}^{\infty}|x_i|^2$ y $b = \sum_{i=1}^{\infty}|y_i|^2$. Si $N \in \mathbb{N}$, entonces los vectores $x = (x_1, \dots, x_N)$ e $y = (y_1, \dots, y_N)$ pertenecen a $\Bbbk^N$ y la desigualdad triangular de la norma asociada al producto interno estándar de $\Bbbk^N$ implica que

$$\begin{aligned}\sum_{i=1}^{N}|x_i + y_i|^2 = \|x + y\|^2 \leq \left(\|x\| + \|y\|\right)^2 = \left(\sqrt{\sum_{i=1}^{N}|x_i|^2} + \sqrt{\sum_{i=1}^{N}|y_i|^2}\right)^2 \\ \leq \left(\sqrt{\sum_{i=1}^{\infty}|x_i|^2} + \sqrt{\sum_{i=1}^{\infty}|y_i|^2}\right)^2 = \left(\sqrt{a} + \sqrt{b}\right)^2.\end{aligned}$$

Esto nos dice que las sumas parciales de la serie $\sum_{i=1}^{\infty}|x_i + y_i|^2$, que tiene términos no negativos, están acotadas y, por lo tanto, que esta serie converge: vemos así que $\xi + \zeta$ es un elemento de ℓ_2, como queríamos.

Mostremos ahora que

$$\text{si } \xi = (xi)_{i\geq 1},\ \zeta = (y_i)_{i\geq 1} \in \ell_2, \text{ entonces la serie } \sum_{i=1}^{\infty} x_i\overline{y_i} \text{ converge absolutamente.} \tag{11}$$

Sean para ello $\xi = (xi)_{i\geq 1}$ y $\zeta = (y_i)_{i\geq 1}$ dos elementos de ℓ_2. De manera similar a lo que hicimos recién, consideremos un entero $N \in \mathbb{N}$ y los vectores $x = (|x_1|, \dots, |x_N|)$ e $y = (|y_1|, \dots, |y_N|)$ de $\Bbbk^N$. La desigualdad de Cauchy–Bunyakovsky–Schwartz para el producto interno estándar de $\Bbbk^N$ nos dice que

$$\sum_{i=1}^{N}|x_i\overline{y_i}| = \left|\langle x, y\rangle_{\Bbbk^N}\right| \leq \|x\|_{\Bbbk^N} \cdot \|y\|_{\Bbbk^N} = \left(\sum_{i=1}^{N}|x_i|\right)^{1/2}\left(\sum_{i=1}^{N}|y_i|\right)^{1/2} \leq \left(\sum_{i=1}^{\infty}|x_i|\right)^{1/2}\left(\sum_{i=1}^{\infty}|y_i|\right)^{1/2}$$

y por lo tanto las sumas parciales de la serie $\sum_{i=1}^{\infty}|x_i\overline{y_i}|$ están acotadas. Esto implica que esa serie converge y que la serie $\sum_{i=1}^{\infty} x_i\overline{y_i}$ converge absolutamente, como afirma (11).

Ahora, en vista de (11), hay una función $\langle -, -\rangle : \ell_2 \times \ell_2 \to \Bbbk$ tal que

$$\langle \xi, \zeta\rangle = \sum_{i=1}^{\infty} x_i\overline{y_i}$$

cada vez que $\xi = (xi)_{i\geq 1}$ y $\zeta = (y_i)_{i\geq 1}$ están en ℓ_2, y es fácil verificar que se trata de un producto interno sobre ℓ_2.

Sea S el subconjunto de ℓ_2 de las sucesiones $\xi = (x_i)_{i\geq 1}$ que tienen un número finito de componentes no nulas; se trata, de hecho, de un subespacio. Afirmamos que $S^\perp = 0$. En efecto, supongamos que $\zeta = (y_i)_{i\geq 1}$ es un elemento de $S^\perp$: para cada $j \in \mathbb{N}$, la sucesión $e = (e_i)_{i\geq 1}$ tal que $e_j = 1$ y $e_i = 0$ si $i \neq j$ es claramente un elemento de S y se tiene que $y_i = \langle \zeta, e\rangle = 0$.

Por supuesto, de esto se sigue que $S^{\perp\perp} = \ell_2$ y, en consecuencia, que $S \subsetneq S^{\perp\perp}$. ◇

§5. Proyectores ortogonales

7.5.1. Si V es un espacio vectorial, decimos que un endomorfismo $p : V \to V$ de V es un ***proyector*** si $p^2 = p$.

Lema. *Sea V un espacio vectorial. Si $p : V \to V$ es un proyector, entonces*

(i) un vector $x \in V$ está en la imagen de p si y solamente si $x = p(x)$, y

(ii) hay una descomposición $V = \mathsf{Im}(p) \oplus \mathsf{Nu}(p)$.

Demostración. Sea $p : V \to V$ un proyector.

(*i*) Si x está en la imagen de p, de manera que existe $y \in V$ tal que $p(y) = x$, entonces $p(x) = p^2(x) = p(y) = x$. Recíprocamente, es evidente que si $p(x) = x$ entonces x está en $\mathsf{Im}(p)$.

(*ii*) Si $x \in V$, entonces $p(x - p(x)) = p(x) - p^2(x) = 0$, así que $x - p(x) \in \mathsf{Nu}(p)$ y

$$x = p(x) + (x - p(x)) \in \mathsf{Im}(p) + \mathsf{Nu}(p).$$

Vemos así que $V = \mathsf{Im}(p) + \mathsf{Nu}(p)$. Esta suma es directa: si $x \in \mathsf{Im}(p) \cap \mathsf{Nu}(p)$, entonces como $x \in \mathsf{Im}(p)$ la primera parte nos dice que $x = p(x)$ y, como $x \in \mathsf{Nu}(p)$, esto implica que $x = 0$. □

7.5.2. Proposición. *Sea V un espacio vectorial y sea S un subespacio de V.*

(i) *Si T es un complemento de S en V, entonces hay un único proyector $p : V \to V$ tal que $\mathsf{Im}(p) = S$ y $\mathsf{Nu}(p) = T$.*

(ii) *Hay un proyector $p : V \to V$ que tiene a S por imagen si y solamente si S posee un complemento T en V.*

Demostración. (i) Supongamos que T es un complemento de S en V, de manera que $V = S \oplus T$. Hay una función $p : V \to V$ tal que si $x \in V$ y $s \in S$ y $t \in T$ son tales que $x = s + t$ entonces $p(x) = s$: esto es consecuencia de que la suma es directa, con lo que esos vectores s y t están unívocamente determinados por x. Mostremos que p es un proyector de V que satisface las condiciones del enunciado:

- Sean x e y vectores de V y $a, b \in \Bbbk$. Si $s_1, s_2 \in S$ y $t_1, t_2 \in T$ son tales que $x = s_1 + t_1$ e $y = s_2 + t_2$, entonces $p(x) = s_1$, $p(y) = s_2$ y, como

$$ax + by = (as_1 + bs_2) + (at_1 + bt_2)$$

con $as_1 + bs_2 \in S$ y $at_1 + bt_s \in T$,

$$p(ax + by) = as_1 + bs_2 = ap(x) + bp(y).$$

Esto nos dice que p es una función lineal.

- Si $x \in V$ y los vectores $s \in S$ y $t \in T$ son tales que $x = s + t$, entonces $p(x) = s \in S$: esto nos dice que $\mathsf{Im}(p) \subseteq S$. Por otro lado, si $s \in S$, entonces es claro que $p(s) = s$ y, por lo tanto, que $S \subseteq \mathsf{Im}(p)$. Vemos así que $\mathsf{Im}(p) = S$ y que $p^2 = p$, de manera que p es un proyector.
- Por otro lado, si $x \in V$ es un elemento de $\mathsf{Nu}(p)$ y $s \in S$ y $t \in T$ son tales que $x = s + t$, entonces $0 = p(x) = s$: esto implica que $x = t \in T$ y, en consecuencia, que $\mathsf{Nu}(p) \subseteq T$. Recíprocamente, si $t \in T$ entonces de la definición de p se siguen inmediatamente que $p(t) = 0$. En definitiva, tenemos que $\mathsf{Nu}(p) = T$.

Supongamos ahora que $q : V \to V$ es otro proyector tal que $\mathsf{Im}(q) = S$ y $\mathsf{Nu}(q) = T$. Sea $x \in V$ y sean $s \in S$ y $t \in T$ tales que $x = s + t$. Como $s \in \mathsf{Im}(q)$, de la Proposición 7.5.1*(i)* sabemos que $s = q(s)$; por otro lado, como $t \in \mathsf{Nu}(q)$, tenemos que $q(t) = 0$. Se sigue de esto, entonces, que $q(x) = q(s) + q(t) = s = p(s)$. Así, es $q = p$ y esto prueba la unicidad que afirma la proposición.

(ii) Si hay un proyector $p : V \to V$ que tiene a S por imagen, entonces de la Proposición 7.5.1*(ii)* se sigue que el núcleo $\mathsf{Nu}(p)$ es un complemento para S. Recíprocamente, la parte (i) que acabamos de probar nos dice que si S posee un complemento, entonces existe un proyector $p : V \to V$ que tiene a S por imagen. □

7.5.3. Sea V un espacio vectorial con producto interno. Si $p : V \to V$ es un proyector tal que $\mathsf{Im}(p) \perp \mathsf{Nu}(p)$, entonces decimos que p es un ***proyector ortogonal***.

Proposición. *Sea V un espacio vectorial con producto interno.*

(i) *Si $p : V \to V$ es un proyector ortogonal, entonces $\mathsf{Im}(p)^\perp = \mathsf{Nu}(p)$ y $\mathsf{Nu}(p)^\perp = \mathsf{Im}(p)$.*

(ii) *Si S es un subespacio de V y existe un proyector ortogonal $p : V \to V$ cuya imagen es S, entonces $S^{\perp\perp} = S$ y $V = S \oplus S^\perp$.*

Demostración. (i) Como p es un proyector ortogonal, tenemos que $\mathsf{Im}(p) \perp \mathsf{Nu}(p)$ y entonces $\mathsf{Im}(p) \subseteq \mathsf{Nu}(p)^\perp$ y $\mathsf{Nu}(p) \subseteq \mathsf{Im}(p)^\perp$. Veamos que valen las igualdades.

Como p es un proyector, sabemos que $V = \mathsf{Im}(p) \oplus \mathsf{Nu}(p)$. Sea $x \in \mathsf{Nu}(p)^\perp$ y sean $y \in \mathsf{Im}(p)$ y $z \in \mathsf{Nu}(p)$ tales que $x = y + z$. Entonces

$$0 = \langle x, z\rangle = \langle y + z, z\rangle = \langle y, z\rangle + \langle z, z\rangle$$

y el primer sumando se anula porque $\mathsf{Im}(p) \perp \mathsf{Nu}(p)$. Así, es $\langle z, z\rangle = 0$, de manera que $z = 0$ y $x = y \in \mathsf{Im}(p)$. Concluimos de esta forma que $\mathsf{Nu}(p)^\perp \subseteq \mathsf{Im}(p)$. De manera enteramente similar puede verse que $\mathsf{Im}(p)^\perp \subseteq \mathsf{Nu}(p)$.

(ii) Si $p : V \to V$ es un proyector ortogonal con $S = \mathsf{Im}(p)$, la primera parte de la proposición nos dice que $S^\perp = \mathsf{Im}(p)^\perp = \mathsf{Nu}(p)$ y que entonces $S^{\perp\perp} = \mathsf{Nu}(p)^\perp = \mathsf{Im}(p) = S$. Por otro lado, como p es un proyector tenemos que $V = \mathsf{Im}(p) \oplus \mathsf{Nu}(p)$, y todo esto prueba lo que queremos. □

7.5.4. En vista de la Proposición 7.5.3*(ii)* y el Ejemplo 7.4.6, no todo subespacio de un espacio vectorial con producto interno es la imagen de un proyector ortogonal. La siguiente proposición nos dice que eso puede ocurrir solamente si el subespacio tiene dimensión infinita:

Proposición. *Sea V un espacio vectorial con producto interno. Si S es un subespacio de V de dimensión finita, entonces $V = S \oplus S^\perp$ y existe exactamente un proyector ortogonal $p : V \to V$ tal que $\mathsf{Im}(p) = S$. Más aún, vale que $\mathsf{Nu}(p) = S^\perp$ y, si $\mathscr{B} = \{x_1, \dots, x_n\}$ es una base ortonormal de S con n elementos, que para cada $x \in V$ se tiene que*

$$p(x) = \sum_{i=1}^{n} \langle x, x_i\rangle x_i.$$

Demostración. Sea S un subespacio de V de dimensión finita n y sea $\mathscr{B} = \{x_1, \dots, x_n\}$ una base de S. La función $p : V \to V$ tal que $p(x) = \sum_{i=1}^{n}\langle x, x_i\rangle x_i$ para todo $x \in V$ es claramente linear y, de acuerdo a la Proposición 7.3.12, se tiene que $x - p(x) \in S^\perp$ para todo $x \in V$.

Si $x \in \mathsf{Nu}(p)$, entonces $x = x - p(x) \in S^\perp$ y, recíprocamente, si $x \in S^\perp$, de manera que en particular $\langle x, x_i\rangle = 0$ para cada $i \in [\![n]\!]$, entonces $p(x) = \sum_{i=1}^{n}\langle x, x_i\rangle x_i = 0$. Así, es $\mathsf{Nu}(p) = S^\perp$. De la definición de p es claro que $\mathsf{Im}(p) \subseteq S$. Por otro lado, si $x \in S$, entonces el Corolario 7.3.6 nos dice que $x = \sum_{i=1}^{n}\langle x, x_i\rangle x_i$ porque $\mathscr{B}$ es una base ortonormal de S y, como consecuencia de esto, que $x = p(x) \in \mathsf{Im}(p)$. Luego $\mathsf{Im}(p) = S$ y $p^2 = p$, como queríamos. □

7.5.5. Hay una relación muy estrecha entre los proyectores ortogonales de un espacio vectorial con producto interno y sus subespacios. Un ejemplo de esto es el siguiente resultado:

Proposición. *Sea V un espacio vectorial con producto interno y sean p, $q : V \to V$ dos proyectores ortogonales.*

(i) *Es $p \circ q = 0$ si y solamente si $\mathsf{Im}(p) \perp \mathsf{Im}(q)$.*

(ii) *Es $p \circ q = p$ si y solamente si $q \circ p = p$ y esto ocurre si y solamente si $\mathsf{Im}(p) \subseteq \mathsf{Im}(q)$.*

(iii) *Es $p \circ q = p \circ p$ si y solamente $p \circ q$ es un proyector ortogonal, y en ese caso se tiene que $\mathsf{Im}(p \circ q) = \mathsf{Im}(p) \cap \mathsf{Im}(q)$.*

Demostración. Observemos primero que

$$\text{si } p : V \to V \text{ es un proyector ortogonal y } x,\ y \in V, \text{ entonces } \langle p(x), y\rangle = \langle x, p(y)\rangle. \tag{12}$$

En efecto, si $x = s + t$ e $y = s' + t'$ con $s, s' \in \mathsf{Im}(p)$ y $t, t' \in \mathsf{Nu}(p)$, entonces

$$\langle p(x), y\rangle = \langle s, s' + t'\rangle = \langle s, s'\rangle = \langle s + t, s'\rangle = \langle x, p(y)\rangle.$$

(i) Supongamos que $p \circ q = 0$. Si $x \in \mathsf{Im}(p)$ e $y \in \mathsf{Im}(q)$, entonces $q(y) = y$ y

$$p(y) = p(q(y)) = 0,$$

de manera que $y \in \mathsf{Nu}(p) = \mathsf{Im}(p)^{\perp}$: esto implica que $x \perp y$ y muestra que $\mathsf{Im}(p) \perp \mathsf{Im}(q)$.

Para probar la implicación recíproca, supongamos que $\mathsf{Im}(p) \perp \mathsf{Im}(q)$ y sea $x \in V$. Como $q(x) \in \mathsf{Im}(q) \subseteq \mathsf{Im}(p)^{\perp} = \mathsf{Nu}(p)$, tenemos que $p(q(x)) = 0$. Vemos así que $p \circ q = 0$.

(ii) Probemos la equivalencia de las tres condiciones del enunciado.

- Supongamos que $p \circ q = p$. Si $x \in V$, entonces usando (12) vemos que para todo $y \in V$ vale que

  $$\langle q(p(x)), y\rangle = \langle p(x), q(y)\rangle = \langle x, p(q(y))\rangle = \langle x, p(y)\rangle = \langle p(x), y\rangle$$

  y, por lo tanto, que $q(p(x)) = p(x)$: esto nos dice que $q \circ p = p$.
- De manera similar, si $q \circ p = p$, entonces para cada x e y en V se tiene que

  $$\langle p(q(x)), y\rangle = \langle q(x), p(y)\rangle = \langle x, q(p(y))\rangle = \langle x, p(y)\rangle = \langle p(x), y\rangle,$$

  así que $p(q(x)) = p(x)$ y, en definitiva, $p \circ q = p$.
- Si $q \circ p = p$ entonces claramente $\mathsf{Im}(p) \subseteq \mathsf{Im}(q)$.
- Recíprocamente, si $\mathsf{Im}(p) \subseteq \mathsf{Im}(q)$, entonces para cada $x \in V$ es $p(x) \in \mathsf{Im}(q)$, así que $q(p(x)) = p(x)$: esto significa quer $q \circ p = p$.

(iii) Supongamos primero que $p \circ q = q \circ p$. Tenemos entonces que

$$(p \circ q)^2 = p \circ q \circ p \circ p = p \circ p \circ q \circ q = p \circ q,$$

así que $r = p \circ q$ es un proyector. Como $\mathsf{Im}(r) = \mathsf{Im}(p \circ q) \subseteq \mathsf{Im}(p)$ e $\mathsf{Im}(r) = \mathsf{Im}(q \circ p) \subseteq \mathsf{Im}(q)$, es claro que $\mathsf{Im}(r) \subseteq \mathsf{Im}(p) \cap \mathsf{Im}(q)$. Recíprocamente, si $x \in \mathsf{Im}(p) \cap \mathsf{Im}(q)$, entonces sabemos que $p(x) = x$ y que $q(x) = x$, así que $x = p(x) = p(q(x)) = r(x) \in \mathsf{Im}(r)$. Esto prueba que $\mathsf{Im}(r) = \mathsf{Im}(p) \cap \mathsf{Im}(q)$.

Supongamos ahora que la composición $r = p \circ q$ es un proyector ortogonal. Si $x \in V$, entonces para cada $y \in V$ vale que

$$\langle q(p(x)), y\rangle = \langle p(x), q(y)\rangle = \langle x, p(q(y))\rangle = \langle x, r(y)\rangle = \langle r(x), y\rangle = \langle p(q(x)), y\rangle$$

y esto nos dice que $q \circ p = p \circ q$. Esto muestra que la condición del enunciado es suficiente. □

§6. El teorema de representación de Riesz

7.6.1. Lema. *Sea V un espacio vectorial con producto interno y para cada $y \in V$ sea la función*

$$\phi_y : x \in V \mapsto \langle x, y \rangle \in \Bbbk.$$

(*i*) *Para cada $y \in V$, la función ϕ_y es lineal, de manera que $\phi_y \in V^*$.*

(*ii*) *La función $\Phi : y \in V \mapsto \phi_y \in V^*$ es sesquilineal, de manera que para cada $y, y' \in V$ y cada $a \in \Bbbk$ se tiene que*

$$\phi_{y+y'} = \phi_y + \phi_{y'}, \qquad \phi_{day} = \overline{a}\phi_y,$$

e inyectiva.

Demostración. (*i*) Sea $y \in V$. Si $x, x' \in V$ y $a, b \in \Bbbk$, entonces

$$\phi_y(ax + bx') = \langle ax + bx', y \rangle = a\langle x, y \rangle + b\langle x', y \rangle = a\phi_y(x) + b\phi_y(x'),$$

así que la función ϕ_y es lineal.

(*ii*) Sean $y, y' \in V$. Para cada $y \in V$, es

$$\phi_{y+y'}(x) = \langle x, y + y' \rangle = \langle x, y \rangle + \langle x, y' \rangle = \phi_y(x) + \phi_{y'}(x) = (\phi_y + \phi_{y'})(x),$$

así que $\phi_{y+y'} = \phi_y + \phi_{y'}$. De manera similar, si $y \in V$ y $a \in \Bbbk$, para cada $x \in V$ se tiene que

$$\phi_{ay}(x) = \langle y, ay \rangle = \overline{a}\langle x, y \rangle = \overline{a}\phi_y(x) = (\overline{a}\phi_y)(x),$$

de manera que $\phi_{ay} = \overline{a}\phi_y$. Esto prueba que la función Φ es sesquilineal.

Finalmente, si $y, y' \in V$ son tales que $\Phi(y) = \Phi(y')$, entonces para cada $x \in V$ se tiene que

$$\langle x, y - y' \rangle = \langle x, y \rangle - \langle x, y' \rangle = \Phi(y)(x) - \Phi(y')(x) = 0,$$

y entonces $y - y' = 0$, esto es, $y = y'$. Vemos así que la función Φ es inyectiva. □

7.6.2. El siguiente resultado es el *teorema de representación de Riesz*, por Frigyes Riesz (1880–1956, Hungría), quien probó un resultado más general para espacios de Hilbert:

Teorema. *Sea V un espacio vectorial con producto interno y de dimensión finita. Si $f \in V^*$, entonces existe un único vector $x_f \in V$ tal que $\phi_{x_f} = x$, esto es, tal que para todo $x \in V$ se tiene que*

$$f(x) = \langle x, x_f \rangle. \tag{13}$$

Demostración. Sea n la dimensión de V, sea $\mathscr{B} = \{x_1, \dots, x_n\}$ una base ortonormal de V y sea

$$x_f = \sum_{i=1}^{n} \overline{f(x_i)} x_i.$$

Si $x \in V$, entonces $x = \sum_{i=1}^{n} \langle x, x_i \rangle x_i$ y

$$f(x) = f\left(\sum_{i=1}^{n} \langle x, x_i \rangle x_i\right) = \sum_{i=1}^{n} \langle x, x_i \rangle f(x_i) = \left\langle x, \sum_{i=1}^{n} \overline{f(x_i)} x_i \right\rangle = \langle x, x_f \rangle.$$

Esto nos dice que x_f satisface la condición (13).

Por otro lado, si $x'_f \in V$ es *otro* vector que la satisface, entonces para todo $x \in V$ se tiene que

$$\langle x, x_f - x'_f \rangle = \langle x, x_f \rangle - \langle x, x'_f \rangle = f(x) - f(x) = 0.$$

Esto es solo posible si $x_f = x'_f$. □

7.6.3. Corolario. *La función* $\Phi : V \to V^*$ *del Lema 7.6.1(ii) es una biyección.*

Demostración. En efecto, el Teorema 7.6.2 nos dice que es sobreyectiva y ese lema nos dice que es inyectiva. □

7.6.4. Ejemplo. En general, si V es un espacio vectorial con producto interno y $f : V \to \mathbb{k}$ es lineal, no existe un vector $y \in V$ tal que $f(x) = \langle x, y \rangle$ para todo $x \in V$. Sea, por ejemplo, $\mathbb{k}^\infty$ el espacio vectorial de las sucesiones $(x_i)_{i \geq 1}$ de elementos de $\mathbb{k}$ con casi todas las componentes nulas, dotado del producto interno tal que si $x = (x_i)_{i \geq 1}$ e $y = (y_i)_{i \geq 1}$ es

$$\langle x, y \rangle = \sum_{i \geq 1} x_i \overline{y_i}.$$

Notemos que esto tiene sentido, ya que cualesquiera sean x e y en V la suma es finita. Para cada $n \geq 1$ sea e_n el elemento de V que tiene todas sus componentes nulas salvo la n-ésima, que es igual a 1, y consideremos la función lineal $f : V \to \mathbb{k}$ tal que $f(x) = \sum_{i \geq 1} x_i$ si $x = (x_i)_{i \geq 1} \in V$. Si hubiera un vector $y = (y_i)_{i \geq 1} \in V$ tal que $f(x) = \langle x, y \rangle$ para todo $x \in V$, tendríamos en particular que para todo $n \geq 1$ sería $\overline{y_n} = \langle e_n, y \rangle = f(e_n) = 1$, lo que es absurdo. ◇

§7. Funciones adjuntas

7.7.1. Sean V y W espacios vectoriales con producto interno y sea $f : V \to W$ una función lineal. Decimos que una función lineal $g : W \to V$ es ***adjunta*** a f si para todo $x \in V$ y todo $y \in W$ se tiene que

$$\langle f(x), y \rangle_W = \langle x, g(y) \rangle_V.$$

Lema. *Sean V y W espacios vectoriales con producto interno. Una función lineal $f : V \to W$ posee a lo sumo* una *función lineal adjunta.*

En vista de este lema, si una función lineal $f : V \to W$ tiene una función adjunta, tiene una sola: podemos en ese caso denotarla sin ambigüedad f^*.

Demostración. Supongamos que $g, g' : W \to V$ son funciones lineales adjuntas a f. Si $y \in W$, entonces para todo $x \in W$ se tiene que

$$\langle x, g(y) - g'(y) \rangle_V = \langle x, g(y) \rangle_V - \langle x, g'(y) \rangle_V = \langle f(x), y \rangle_W - \langle f(x), y \rangle_W = 0,$$

así que $g(y) = g'(y)$. Esto nos dice que $g = g'$. □

7.7.2. Ejemplos.

(a) Si V es un espacio vectorial con producto interno y $\lambda \in \Bbbk$ es un escalar, entonces la función lineal $\lambda \mathrm{id}_V : V \to V$ posee una adjunta, que es $\overline{\lambda}\mathrm{id}_V : V \to V$.

(b) Sea $V = C_c^\infty(\mathbb{R})$ el espacio vectorial de todas las funciones infinitamente diferenciables $\mathbb{R} \to \Bbbk$ que tienen soporte compacto, dotado del producto interno $\langle -, - \rangle$ tal que

$$\langle f, g \rangle = \int_{-\infty}^{\infty} f(x)\overline{g(x)} \mathrm{d}x$$

para cada f, $g \in V$; notemos que esta integral tiene sentido, porque el soporte de $f\overline{g}$ es compacto y la función es allí continua. Si $f \in V$, entonces también $f' \in V$, así que podemos considerar la función lineal

$$D : f \in V \mapsto f' \in V.$$

Si $f, g \in V$, entonces la fórmula de integración por partes nos dice que

$$\langle Df, g \rangle = \int_{-\infty}^{\infty} f'(x)\overline{g(x)} \mathrm{d}x = -\int_{-\infty}^{\infty} f(x)\overline{g'(x)} \mathrm{d}x = \langle f, -Dg \rangle,$$

y esto significa que D posee una función adjunta y que $D^* = -D$. ◇

7.7.3. Ejemplo. En general, un endomorfismo de un espacio vectorial con producto interno no tiene una función adjunta. Para dar un ejemplo, sea $V = \Bbbk^\infty$ el espacio vectorial de las sucesiones $(x_i)_{i \geq 1}$ de elementos de $\Bbbk$ con casi todas las componentes nulas, con el producto interno tal que si $x = (x_i)_{i \geq 1}$ e $y = (y_i)_{i \geq 1}$ es

$$\langle x, y \rangle = \sum_{i \geq 1} x_i \overline{y_i}.$$

Notemos que esto tiene sentido, ya que cualesquiera sean x e y en V la suma es finita. Para cada $n \in \mathbb{N}$ escribimos e_n al elemento de V que tiene todas sus componentes nulas salvo la n-ésima, que es igual a 1, y consideramos la función lineal $L : V \to \Bbbk$ tal que si $x = (x_i)_{i \geq 1}$ es un elemento de V entonces

$$L(x) = \sum_{i \geq 1} x_i.$$

Supongamos que L posee una función adjunta $L^* : \Bbbk \to V$, de manera que, en particular, si escribimos $L^*(1) = (u_i)_{i \geq 1}$, es

$$1 = \langle L(e_n), 1 \rangle = \langle e_n, L^*(1) \rangle = \overline{u_i}$$

para cada $i \geq 1$: esto es absurdo, porque nos dice que *todas* las componentes de $L^*(1) \in \Bbbk^\infty$ son no nulas. Este ejemplo está en evidente relación con el Ejemplo 7.6.4 visto arriba. ◇

7.7.4. Proposición. *Sean V y W dos espacios vectoriales con producto interno.*

(i) *Si f, $g : V \to W$ son funciones lineales que poseen funciones adjuntas, entonces para cada a, $b \in \Bbbk$ la función lineal $af + bg : V \to W$ posee función adjunta y, de hecho, se tiene que*

$$(af + bg)^* = \overline{a}\, f^* + \overline{b}\, g^*.$$

(ii) *Si $f : V \to W$ es una función lineal que posee una función adjunta, entonces f^* también posee una función adjunta y*

$$(f^*)^* = f.$$

Demostración. (i) Si $x \in V$ e $y \in W$, entonces

$$\begin{aligned}\langle (af + bg)(x), y\rangle &= \langle af(x) + bf(x), y\rangle = a\langle f(x), y\rangle + b\langle g(x), y\rangle \\ &= a\langle x, f^*(y)\rangle + b\langle x, g^*(y)\rangle = \langle x, \overline{a}\, f^*(y) + \overline{b}\, g^*(y)\rangle \\ &= \langle x, (\overline{a}\, f^* + \overline{b}\, g^*)(y)\rangle.\end{aligned}$$

Esto significa que la función lineal $\overline{a}\, f^* + \overline{b}\, g^* : V \to W$ es adjunta a $af + gb$.

(ii) Si $x \in V$ e $y \in W$, tenemos que

$$\langle f^*(x), y\rangle = \overline{\langle y, f^*(x)\rangle} = \overline{\langle f(y), x\rangle} = \langle x, f(y)\rangle,$$

y esto nos dice que $f : V \to W$ es adjunta a $f^* : W \to V$, esto es, que $(f^*)^* = f$. □

7.7.5. Proposición.

(i) *Si V es un espacio con producto interno, entonces la función identidad $\mathsf{id}_V : V \to V$ posee función adjunta y*

$$(\mathsf{id}_V)^* = \mathsf{id}_V.$$

(ii) *Si V, W y U son espacios vectoriales con producto interno y $f : V \to W$ y $g : W \to U$ son funciones lineales que poseen funciones adjuntas, entonces la composición $g \circ f : V \to U$ posee función adjunta y*

$$(g \circ f)^* = f^* \circ g^*.$$

Demostración. Para ver la primera parte, es suficiente notar que si x e y son elementos de V se tiene que

$$\langle \mathsf{id}_V(x), y\rangle = \langle x, y\rangle = \langle x, \mathsf{id}_V(y)\rangle,$$

así que $(\mathsf{id}_V)^* = \mathsf{id}_V$. Por otro lado, si $f : V \to W$ y $g : W \to U$ son funciones lineales que poseen funciones adjuntas, entonces para cada $x \in V$ y cada $y \in U$ se tiene que

$$\langle (g \circ f)(x), y\rangle = \langle g(f(x)), y\rangle = \langle f(x), g^*(y)\rangle = \langle x, f^*(g^*(y))\rangle = \langle x, (f^* \circ g^*)(y)\rangle.$$

Esto significa que $g \circ f$ tiene como función adjunta a $f^* \circ g^*$. □

7.7.6. Teorema. *Sean V y W espacios vectoriales con producto interno. Si V tiene dimensión finita, entonces Toda función lineal $f : V \to W$ posee una función adjunta $f^* : W \to V$.*

Demostración. Si $y \in V$, la función

$$\psi_y : x \in V \mapsto \langle f(x), y\rangle_W \in \Bbbk$$

es lineal, así que el Teorema 7.6.2 nos dice que existe un único un vector $f^*(y) \in V$ con la propiedad de que para todo $x \in V$ es $\psi_y(x) = \langle x, f^*(y)\rangle_V$, esto es, para todo $y \in V$ es

$$\langle f(x), y\rangle_V = \langle x, f^*(y)\rangle_W.$$

De esta forma tenemos definida una función $f^* : W \to V$ que satisface la identidad deseada. Para terminar la prueba del teorema resta únicamente mostrar que f^* es lineal.

Sean $y, y' \in W$. Para todo $x \in V$ se tiene que

$$\begin{aligned}\langle x, f^*(y+y')\rangle_V &= \langle f(x), y+y'\rangle_W = \langle f(x), y\rangle_W + \langle f(x), y'\rangle_W \\ &= \langle x, f^*(y)\rangle_W + \langle x, f^*(y')\rangle_W = \langle x, f^*(y) + f^*(y)\rangle_W,\end{aligned}$$

así que $f^*(y+y') = f^*(y) + f^*(y')$. Por otro lado, si $y \in W$ y $\lambda \in \Bbbk$, para cada $x \in W$ tenemos que

$$\begin{aligned}\langle x, f^*(\lambda y)\rangle_V &= \langle f(x), \lambda y\rangle_W = \langle \overline{\lambda} f(x), y\rangle_W = \langle f(\overline{\lambda} x), y\rangle_W = \langle \overline{\lambda} x, f^*(y)\rangle_V \\ &= \langle x, \lambda f^*(y)\rangle_V,\end{aligned}$$

de manera que $f^*(\lambda w) = \lambda f^*(w)$. Concluimos que f^* es lineal, como queríamos. □

7.7.7. Proposición. *Sean V y W espacios vectoriales con producto interno de dimensiones finita y positivas m y n y sean $\mathscr{B}$ y $\mathscr{B}'$ bases ordenadas ortonormales de V y de W, respectivamente. Si $f : V \to W$ es una función lineal, entonces la matriz de la función adjunta $f^* : W \to V$ con respecto a las bases $\mathscr{B}'$ y $\mathscr{B}$ es*

$$[f^*]^{\mathscr{B}'}_{\mathscr{B}} = \overline{([f]^{\mathscr{B}}_{\mathscr{B}'})^{\mathsf{t}}}.$$

Demostración. Supongamos que las bases de V y de W son $\mathscr{B} = (x_1, \dots, x_m)$ y $\mathscr{B}' = (y_1, \dots, y_n)$, y que $[f]^{\mathscr{B}}_{\mathscr{B}'} = (a_{i,j}) \in \mathsf{M}_{n,m}(\Bbbk)$. Si $i \in [\![m]\!]$ y $j \in [\![n]\!]$, es

$$\overline{\langle f^*(y_j), x_i\rangle} = \langle x_i, f^*(y_j)\rangle = \langle f(x_i), y_j\rangle = \left\langle \sum_{k=1}^m a_{k,i} y_k, y_j \right\rangle = \sum_{k=1}^m a_{k,i} \langle y_k, y_j\rangle = a_{j,i}$$

porque la base $\mathscr{B}'$ es ortonormal y entonces $\langle f^*(y_j), x_i\rangle = \overline{a_{j,i}}$. De acuerdo al Corolario 7.3.6, se sigue de esto que

$$f^*(y_j) = \sum_{i=1}^m \overline{a_{j,i}}\, x_i$$

y entonces la matriz de f^* con respecto a las bases $\mathscr{B}'$ y $\mathscr{B}$ es $(\overline{a_{j,i}})_{i,j}$, esto es, la matriz que se obtiene de $[f]^{\mathscr{B}}_{\mathscr{B}'}$ transponiendo y conjugando, como afirma la proposición. □

7.7.8. Ejemplo. Es importante observar que la igualdad que afirma la Proposición 7.7.7 no es cierta, en general, si las bases consideradas no son ortonormales. Demos un ejemplo de esto.

Consideremos el espacio vectorial $\Bbbk^2$ dotado de su producto interno estándar y la matriz $A = \left(\begin{smallmatrix} 0 & 1 \\ 0 & 0 \end{smallmatrix}\right) \in \mathsf{M}_2(\Bbbk)$, y sea $f : x \in \Bbbk^2 \mapsto Ax \in \Bbbk^2$. Si $\mathscr{B} = (e_1, e_2)$ es la base ordenada estándar de $\Bbbk^2$, entonces sabemos que $[f]^{\mathscr{B}}_{\mathscr{B}} = A$ y, como $\mathscr{B}$ es de hecho una base ortonormal, que $[f^*]^{\mathscr{B}}_{\mathscr{B}} = A^{\mathrm{t}}$, la matriz transpuesta de A. En cambio, si ponemos $e_2' = e_1 + e_2$, entonces $\mathscr{B}' = (e_1, e_2')$ también es una base ordenada de $\Bbbk^2$, se tiene que $[f]^{\mathscr{B}'}_{\mathscr{B}'} = \left(\begin{smallmatrix} 0 & 1 \\ 0 & 0 \end{smallmatrix}\right)$ mientras que $[f^*]^{\mathscr{B}'}_{\mathscr{B}'} = \left(\begin{smallmatrix} -1 & 0 \\ 1 & 0 \end{smallmatrix}\right)$. ◇

7.7.9. Proposición. *Sean V y W espacios vectoriales con producto interno y sea $f : V \to W$ una función lineal que posee adjunta. Entonces*

(*i*) $\mathsf{Nu}(f^*) = \mathsf{Im}(f)^\perp$,

(*ii*) $\mathsf{Im}(f^*) \subseteq \mathsf{Nu}(f)^\perp$, *y*

(*iii*) $\mathsf{Im}(f^*)^\perp \subseteq \mathsf{Nu}(f^*)$.

Si V tiene dimensión finita, entonces se tiene además que

(*iv*) $\mathsf{Im}(f^*) = \mathsf{Nu}(f)^\perp$.

Demostración. (*i*) Sea $x \in \mathsf{Nu}(f^*)$. Si $y \in \mathsf{Im}(f)$, de manera que existe $z \in V$ tal que $f(z) = y$, entonces $\langle y, x\rangle = \langle f(z), x\rangle = \langle z, f^*(x)\rangle = 0$. Esto nos dice que $x \in \mathsf{Im}(f)^\perp$. Recíprocamente, si $x \in \mathsf{Im}(()f)^\perp$, se tiene que para todo $y \in V$ es $\langle y, f^*(x)\rangle = \langle f(y), x\rangle = 0$, así que $f^*(x) = 0$, esto es, $x \in \mathsf{Nu}(f^*)$.

(*ii*) Sea $x \in \mathsf{Im}(f^*)$, de manera que existe $y \in V$ tal que $x = f^*(y)$. Para cada $z \in \mathsf{Nu}(f)$ es $\langle z, x\rangle = \langle z, f^*(y)\rangle = \langle f(z), x\rangle = 0$, así que $x \in \mathsf{Nu}(f)^\perp$: hemos probado que $\mathsf{Im}(f^*) \subseteq \mathsf{Nu}(f)^\perp$.

(*iii*) Sea $x \in \mathsf{Im}(f^*)^\perp$. Si $y \in V$, entonces $\langle f(x), y\rangle = \langle x, f^*(y)\rangle = 0$, así que $f(x) = 0$. Vemos de esta forma que $\mathsf{Im}(f^*)^\perp \subseteq \mathsf{Nu}(f^*)$.

(*iv*) Si V tiene dimensión finita, usando la Proposición 7.4.2*(iii)* y la Proposición 7.4.4 y las partes (*ii*) y (*iii*) de esta proposición, tenemos que $\mathsf{Nu}(f^*)^\perp \subseteq \mathsf{Im}(f^*)^{\perp\perp} = \mathsf{Im}(f^*)$. □

7.7.10. Ejemplo. La igualdad de la parte (*iv*) de la Proposición 7.7.9 en general no vale. Hacer. ◇

§8. Funciones lineales autoadjuntas

7.8.1. Sea V un espacio vectorial con producto interno. Si $f : V \to V$ es una función lineal que posee adjunta f^*, decimos que f es ***autoadjunta*** si $f^* = f$. En otras palabras, la función lineal f es autoadjunta si y solamente si para cada x, $y \in V$ si tiene que

$$\langle f(x), y\rangle = \langle x, f(y)\rangle.$$

7.8.2. Proposición. *Sea V un espacio vectorial con producto interno y de dimensión finita y sea $\mathscr{B}$ una base ortonormal de V. Una función lineal $f : V \to V$ es autoadjunta si y solamente si*

$$[f]^{\mathscr{B}}_{\mathscr{B}} = \overline{([f]^{\mathscr{B}}_{\mathscr{B}})^{\mathrm{t}}}.$$

En particular,

(i) *si $\Bbbk = \mathbb{R}$, f es autoadjunta si y solamente si la matriz $[f]^{\mathscr{B}}_{\mathscr{B}}$ es simétrica, y*
(ii) *si $\Bbbk = \mathbb{C}$, f es autoadjunta si y solamente si la matriz $[f]^{\mathscr{B}}_{\mathscr{B}}$ es hermitiana.*

Demostración. Esto es consecuencia inmediata de la Proposición 7.7.7 y de que dos transformaciones lineales $V \to V$ son iguales si y solamente si sus matrices con respecto a la base $\mathscr{B}$ coinciden. □

7.8.3. Proposición. *Sea V un espacio vectorial con producto interno y sea $p : V \to V$ un proyector. Entonces p es ortogonal si y solamente si es autoadjunto.*

Demostración. Supongamos primero que el proyector p es ortogonal. Sean x, $y \in V$. Como $V = \mathsf{Im}(p) \oplus \mathsf{Nu}(p)$, existen x', $y' \in \mathsf{Im}(p)$ y x'', $y'' \in \mathsf{Nu}(p)$ tales que $x = x' + x''$ e $y = y' + y''$. Más aún, $p(x) = x'$, $p(y) = y'$ y, como por hipótesis es $\mathsf{Im}(p) \perp \mathsf{Nu}(p)$, es $\langle x', y''\rangle = \langle x'', y'\rangle = 0$. Usando todo esto, vemos que

$$\begin{aligned}\langle p(x), y\rangle &= \langle x', y' + y''\rangle = \langle x', y'\rangle + \langle x', y''\rangle = \langle x', y'\rangle = \langle x', y'\rangle + \langle x'', y'\rangle \\ &= \langle x' + x'', y'\rangle = \langle x, p(y)\rangle.\end{aligned}$$

Esto nos dice que p posee adjunta y que, de hecho, $p^* = p$.

Supongamos ahora que p es autoadjunto. Sean $x \in \mathsf{Im}(p)$ e $y \in \mathsf{Nu}(p)$. Como p es autoadjunto,

$$\langle x, y\rangle = \langle p(x), y\rangle = \langle x, p(y)\rangle = 0.$$

Esto nos dice que $x \perp y$ y, en definitiva, que $\mathsf{Im}(p) \perp \mathsf{Nu}(p)$. □

7.8.4. Proposición. *Sea V un espacio vectorial con producto interno y sea $f : V \to V$ una función lineal autoadjunta.*

(i) *Todo autovalor de f es real.*
(ii) *Si x, $y \in V$ son autovectores de f correspondientes a autovalores distintos, entonces $x \perp y$.*

Demostración. (i) Sea $\lambda \in \Bbbk$ un autovalor y sea $x \in V$ un autovector de f de autovalor λ. Entonces

$$\lambda\langle x, x\rangle = \langle \lambda x, x\rangle = \langle f(x), x\rangle = \langle x, f(x)\rangle = \langle x, \lambda x\rangle = \overline{\lambda}\langle x, x\rangle$$

y, como $\langle x, x\rangle \neq 0$, esto implica que $\lambda = \overline{\lambda}$, esto es, que $\lambda \in \mathbb{R}$.

(ii) Supongamos que x, $y \in V$ son autovectores de autovalores λ, $\mu \in \Bbbk$, respectivamente, y que $\lambda \neq \mu$. De la parte (i) sabemos que λ, $\mu \in \mathbb{R}$ y entonces

$$\lambda\langle x, y\rangle = \langle \lambda x, y\rangle = \langle f(x), y\rangle = \langle x, f(y)\rangle = \langle x, \mu y\rangle = \overline{\mu}\langle x, y\rangle = \mu\langle x, y\rangle.$$

Vemos así que $(\lambda - \mu)\langle v, w\rangle = 0$ y, como $\lambda \neq \mu$, que $\langle v, w\rangle = 0$. □

7.8.5. Proposición. *Sea V un espacio vectorial con producto interno y de dimensión finita. Una función lineal $f : V \to V$ que es autoadjunta posee un autovalor.*

Demostración. Si $\Bbbk = \mathbb{C}$, ya sabemos que todo endomorfismo de un espacio vectorial de dimensión finita tiene un autovalor, ya que el cuerpo $\mathbb{C}$ es algebraicamente cerrado. Nos queda entonces solamente considerar el caso en que $\Bbbk = \mathbb{R}$.

Sea $n = \dim V$, sea $\mathscr{B}$ una base ortonormal de V y sea $A = [f]_{\mathscr{B}}^{\mathscr{B}} \in M_n(\mathbb{R})$. Como f es autoadjunta, la matriz A es simétrica. Consideremos el endomorfismo $g : x \in \mathbb{C}^n \mapsto Ax \in \mathbb{C}^n$ del espacio vectorial complejo $\mathbb{C}^n$ y sea $\mathscr{B}'$ la base ordenada estándar de $\mathbb{C}^n$. Si dotamos a $\mathbb{C}^n$ de su producto interno estándar, la base $\mathscr{B}'$ es ortonormal, así que, como $[g]_{\mathscr{B}'}^{\mathscr{B}'} = A$ es una matriz hermitiana (ya que es simétrica y real), la función g es autoadjunta y todos sus autovalores son reales. Esto implica que el polinomio característico χ_g de g tiene todas sus raíces reales. Como este polinomio coincide con el polinomio característico χ_A de A y éste último con el polinomio característico χ_f de f, vemos que χ_f tiene sus raíces en $\mathbb{R}$: esto nos dice que f posee algún autovalor. □

7.8.6. Podemos ahora probar el resultado más importante sobre las funciones lineales autoadjuntas: son diagonalizables.

Proposición. *Sea V un espacio vectorial con producto interno y de dimensión finita. Si $f : V \to V$ es una función lineal autoadjunta, existe una base ortonormal $\mathscr{B}$ cuyos elementos son autovectores de f y, en particular, la matriz $[f]_{\mathscr{B}}^{\mathscr{B}}$ es diagonal.*

Demostración. Hacemos inducción en $\dim V$. Es claro que cuando $\dim V = 0$ no hay nada que probar, así que supongamos que $\dim V > 0$.

Sea $f : V \to V$ una función lineal autoadjunta. De acuerdo a la proposición anterior, existe un autovalor $\lambda \in \Bbbk$ y entonces existe $x_1 \in V \setminus 0$ tal que $f(x_1) = \lambda x_1$. Sin pérdida de generalidad, podemos suponer que $\|x_1\| = 1$; si ése no fuese el caso, podríamos simplemente reemplazar a x_1 por el vector $x_1/\|x_1\|$.

Sea $W = \langle x_1 \rangle^\perp$. Si $x \in W$, entonces

$$\langle f(x), x_1 \rangle = \langle x, f(x_1) \rangle = \langle x, \lambda x_1 \rangle = \lambda \langle x, x_1 \rangle = 0,$$

de manera que $f(x) \in W$: esto nos dice que el subespacio W es f-invariante. Dotemos a W del producto interno que se obtiene restringiendo el de V y sea $f_W : W \to W$ la restricción de f a W. Es inmediato ver que f_W es entonces un endomorfismo autoadjunto de W y, como $\dim W = \dim V - 1$, inductivamente podemos suponer que hay una base ordenada ortonormal $(x_2, \dots, x_n)$ de W cuyos elementos son autovectores de f_W. Por supuesto, se sigue inmediatamente de esto que $\mathscr{B} = (x_1, \dots, x_n)$ es una base ordenada ortonormal de V cuyos elementos son autovectores de f, lo que prueba la proposición. □

7.8.7. Corolario. *Sea V un espacio vectorial con producto interno y de dimensión finita y sea $f : V \to V$ una función lineal. La función f es autoadjunta si y solamente si existe una base ordenada ortonormal $\mathscr{B}$ de V tal que la matriz $[f]_{\mathscr{B}}^{\mathscr{B}}$ es diagonal y real.*

Demostración. La necesidad de la condición es consecuencia de la Proposición 7.8.6 y de la Proposición 7.8.4*(i)*. La suficiente, por su parte, sigue inmediatamente de la Proposición 7.7.7. □

§9. Funciones lineales normales

7.9.1. Sea V un espacio vectorial con producto interno. Una función lineal $f : V \to V$ es ***normal*** si posee adjunta f^* y

$$f^* f = f f^*.$$

Es claro que si $f : V \to V$ es autoadjunta, de manera que $f^* = f$, entonces f es normal: en ese caso es $f^* f = f^2 = f f^*$. La implicación recíproca es falsa.

7.9.2. Ejemplo. Consideremos a $\mathbb{R}^2$ dotado de su producto interno usual y sea $\mathscr{B}$ la base ordenada estándar de $\mathbb{R}^2$. Sea $\alpha \in \mathbb{R}$ y sea $f : \mathbb{R}^2 \to \mathbb{R}^2$ la función lineal tal que

$$[f]^{\mathscr{B}}_{\mathscr{B}} = \begin{pmatrix} \cos\alpha & \sin\alpha \\ -\sin\alpha & \cos\alpha \end{pmatrix}$$

Como $\mathscr{B}$ es una base ortonormal, sabemos que

$$[f^*]^{\mathscr{B}}_{\mathscr{B}} = \overline{([f]^{\mathscr{B}}_{\mathscr{B}})^{\mathsf{t}}} = \begin{pmatrix} \cos\alpha & -\sin\alpha \\ \sin\alpha & \cos\alpha \end{pmatrix},$$

y un cálculo inmediato muestra que $[f^*]^{\mathscr{B}}_{\mathscr{B}}[f]^{\mathscr{B}}_{\mathscr{B}} = I_2 = [f]^{\mathscr{B}}_{\mathscr{B}}[f^*]^{\mathscr{B}}_{\mathscr{B}}$, la matriz identidad, de manera que $f^* f = f f^* = \mathrm{id}_V$. Así, f es normal cualquiera sea $\alpha \in \mathbb{R}$. Es claro, por otro lado, que $f^* \neq f$ si el número α no es un múltiplo entero de π. ◇

7.9.3. Proposición. *Sea V un espacio vectorial con producto interno. Sea $f : V \to V$ una función lineal normal y sean $x \in V$ y $\lambda \in \Bbbk$. Las siguientes afirmaciones son equivalentes:*

(a) *x es un autovector para f de autovalor λ.*

(b) *x es un autovector para f^* de autovalor $\overline{\lambda}$.*

Demostración. Sea $g = f - \lambda \mathrm{id}_V$. Sabemos que g posee adjunta y que $g^* = f^* - \overline{\lambda}\mathrm{id}_V$ y, como f es normal, es inmediato verificar que g es normal. Usando esto, tenemos que

$$\|g(x)\|^2 = \langle g(x), g(x)\rangle = \langle x, g^*(g(x))\rangle = \langle x, g(g^*(x))\rangle = \langle g^*(x), g^*(x)\rangle = \|g^*(x)\|^2.$$

Esto nos dice que $g(v) = 0$ sii $g^*(v) = 0$, es decir, que $f(v) = \lambda v$ si y solamente si $f^*(v) = \overline{\lambda} v$, que es precisamente lo que afirma la proposición. □

7.9.4. Teorema. *Sea V un espacio vectorial complejo con producto interno y de dimensión finita. Si $f : V \to V$ es una función lineal normal, entonces existe una base ortonormal de V cuyos elementos son autovectores de f.*

Demostración. Hagamos inducción con respecto a la dimensión de V, notando que si $V = 0$ no hay nada que probar. Como $\mathbb{C}$ es algebraicamente cerrado, sabemos que existen $\lambda \in \mathbb{C}$ y $x \in V$ tales que $\|x\| = 1$ y $f(x) = \lambda x$. De la Proposición 7.9.3, entonces, vale también que $f^*(x) = \overline{\lambda} x$.

Sea $W = \langle x \rangle^\perp$. Si $y \in W$, entonces $y \perp x$ y tenemos que

$$\langle f(y), x \rangle = \langle y, f^*(x) \rangle = \langle y, \overline{\lambda} x \rangle = \lambda \langle y, x \rangle = 0$$

y

$$\langle f^*(y), x \rangle = \langle y, f(x) \rangle = \langle y, \lambda x \rangle = \overline{\lambda} \langle y, x \rangle = 0.$$

Vemos así que el subespacio W es f- y f^*-invariante y que, en particular, podemos considerar las restricciones f_W, $(f^*)_W : W \to W$. Es inmediato verificar que f_W posee una adjunta y que, de hecho, $(f_W)^* = (f^*)_W$. Más aún, la función f_W es normal.

Como $\dim W = \dim V - 1$, podemos suponer que el teorema el cierto para f_W y, entonces, que existe una base ortonormal $\mathscr{B}'$ de W formada por autovectores de f_W. Es claro entonces que $\mathscr{B} = \{x\} \cup \mathscr{B}'$ es una base ortonormal de V formada por autovectores de f. □

7.9.5. Corolario. *Sea V un espacio vectorial complejo con producto interno y de dimensión finita y sea $f : V \to V$ una transformación lineal. Las siguientes afirmaciones son equivalentes:*

(a) *La función f es normal.*

(b) *Existe una base ortonormal $\mathscr{B}$ de V formada por autovectores de f.*

Demostración. La implicación $(a) \Rightarrow (b)$ es el contenido del Teorema 7.9.4. Veamos la recíproca. Si $\mathscr{B}$ es una base ortonormal formada por autovectores de f, entonces la matriz $[f]^{\mathscr{B}}_{\mathscr{B}}$ es diagonal y lo mismo es cierto de la matriz $[f^*]^{\mathscr{B}}_{\mathscr{B}} = \overline{([f]^{\mathscr{B}}_{\mathscr{B}})^{\mathsf{t}}}$. Esto implica que estas dos matrices $[f]^{\mathscr{B}}_{\mathscr{B}}$ y $[f^*]^{\mathscr{B}}_{\mathscr{B}}$ conmutan y, por lo tanto, que f y f^* conmutan: en otras palabras, f es normal. □

7.9.6. El siguiente resultado es conocido como el *Teorema espectral para transformaciones normales*:

Proposición. *Sea V un espacio vectorial complejo con producto interno y de dimensión finita y sea $f : V \to V$ una transformación lineal normal. Sean $\lambda_1, \ldots, \lambda_n \in \mathbb{C}$ los autovalores distintos de f y para cada $i \in [\![n]\!]$ sea $V_i = \mathsf{Nu}(f - \lambda_i \mathsf{id}_V)$ y $p_i : V \to V$ el proyector ortogonal de imagen V_i. Entonces:*

(*i*) *Si i, $j \in [\![n]\!]$ son distintos, entonces $V_i \perp V_j$ y $p_i p_j = 0$.*

(*ii*) *Hay una descomposición en suma directa $V = V_1 \oplus \cdots \oplus V_n$.*

(*iii*) *Se tiene que $\mathsf{id}_V = p_1 + \cdots + p_n$ y $f = \lambda_1 p_1 + \cdots + \lambda_n p_n$.*

Demostración. Si i, $j \in [\![n]\!]$ son distintos y $x \in V_i \setminus 0$ e $y \in V_j \setminus 0$, entonces x e y son autovectores de f de autovalores λ_i y λ_j y la Proposición 7.9.3 nos dice que $f^*(y) = \overline{\lambda}_j y$ y que, en consecuencia,

$$\lambda_i \langle x, y \rangle = \langle \lambda_i x, y \rangle = \langle f(x), y \rangle = \langle x, f^*(y) \rangle = \langle x, \overline{\lambda}_j y \rangle = \lambda_j \langle x, y \rangle. \tag{14}$$

Como $\lambda_i \neq \lambda_j$, esto implica que $\langle x, y\rangle = 0$. Concluimos de esta forma que $V_i \perp V_j$ y, usando la Proposición 7.5.5, que $p_i p_j = 0$. Más aún, de acuerdo al Teorema 7.9.4, hay una base de autovectores de V así que, de hecho, es $V = \bigoplus_{i=1}^n V_i$.

Si $x \in V$, entonces existen $x_1 \in V_1, \dots, x_n \in V_n$ tales que $x = x_1 + \cdots + x_n$. Si $i, j \in [\![n]\!]$, se tiene que

$$p_i(x_j) = \begin{cases} x_i, & \text{si } i = j; \\ 0, & \text{en caso contrario.} \end{cases}$$

En efecto, en el primer caso esto es así porque $x_j \in V_i$ y p_i es el proyector ortogonal con imagen V_i, en el segundo porque $x_j \in V_j \subseteq V_i^{\perp} = \mathsf{Nu}(()p_i)$. Como consecuencia de esto se tiene que

$$(p_1 + \cdots + p_n)(x) = \sum_{i=1}^n \sum_{j=1}^n p_i(x_j) = \sum_{i=1}^n p_i(x_i) = x$$

y

$$\begin{aligned}(\lambda_1 p_1 + \cdots + \lambda_n p_n)(x) &= \sum_{i=1}^n \sum_{j=1}^n \lambda_i p_i(x_j) = \sum_{i=1}^n \lambda_i p_i(x_i) = \sum_{i=1}^n \lambda_i x_i \\ &= \sum_{i=1}^n f(x_i) = f\left(\sum_{i=1}^n x_i\right) = f(x)\end{aligned}$$

y esto implica que $p_1 + \cdots + p_n = \mathsf{id}_V$ y que $\lambda_1 p_1 + \cdots + \lambda_n p_n = f$. □

7.9.7. Lema. *Sea V un espacio vectorial, sean $\lambda_1, \dots, \lambda_n \in \Bbbk$ escalares y sean $p_1, \dots, p_n : V \to V$ proyectores ortogonales tales que $\sum_{i=1}^n p_i = \mathsf{id}_V$ y $p_i p_j = 0$ siempre que $i, j \in [\![n]\!]$ son distintos. Si $f = \sum_{i=1}^n \lambda_i p_i$, entonces para cada $h \in \Bbbk[X]$ es*

$$h(f) = \sum_{i=1}^n h(\lambda_i) p_i.$$

Demostración. Por linealidad, es suficiente probar la igualdad del enunciado para el caso particular en que $h = X^r$ es un monomio y hacemos esto procediendo por inducción en r.

Cuando $r = 0$, de manera que $h = 1$, la igualdad (14) vale porque por hipótesis es $\mathsf{id}_V = \sum_{i=1}^n p_i$. Por otro lado, si $p = X^{r+1}$, entonces

$$h(f) = f^{r+1} = f^r f = \left(\sum_{i=1}^n \lambda_i^r p_i\right)\left(\sum_{j=1}^n \lambda_i p_i\right) = \sum_{i=1}^n \sum_{j=1}^n \lambda_i^r \lambda_j p_i p_j$$

y, como $p_i p_j = 0$ si $i \neq j$, esto es

$$= \sum_{i=1}^n \lambda_i^{r+1} p_i = \sum_{i=1}^n h(\lambda_i) p_i.$$

Esto completa la inducción. □

7.9.8. Proposición. *Sea V un espacio vectorial complejo con producto interno y de dimensión finita y sea $f : V \to V$ una función lineal. Las siguientes afirmaciones son equivalentes:*

(a) *La función f es normal.*

(b) *Existe un polinomio $h \in \mathbb{C}[X]$ tal que $f^* = p(f)$.*

Demostración. $(b) \Rightarrow (a)$ Si existe $h \in \mathbb{C}[X]$ tal que $f^* = h(f)$, entonces

$$f^* \circ f = h(f) \circ f = (hX)(f) = (Xh)(f) = f \circ h(f) = f \circ f^*,$$

así que f es normal.

$(a) \Rightarrow (b)$ De acuerdo al Teorema 7.9.6, hay escalares $\lambda_1, \ldots, \lambda_n \in \mathbb{C}$ distintos dos a dos y proyectores ortogonales $p_1, \ldots, p_n : V \to V$ tales que $p_i p_j = 0$ cada vez que $i, j \in [\![n]\!]$ son distintos. Consideremos el polinomio

$$h(X) = \sum_{i=1}^{n} \overline{\lambda}_i \prod_{\substack{1 \leq j \leq n \\ j \neq i}} \frac{X - \lambda_j}{\lambda_i - \lambda_j} \in \mathbb{C}[X].$$

Es inmediato verificar que $h(\lambda_i) = \overline{\lambda}_i$ para cada $i \in [\![n]\!]$ y entonces, de acuerdo al Lema 7.9.7, se tiene que

$$h(f) = \sum_{i=1}^{n} h(\lambda_i) p_i = \sum_{i=1}^{n} \overline{\lambda}_i p_i.$$

Por otro lado, como los proyectores p_i son ortogonales y, entonces, autoadjuntos, tenemos que

$$f^* = \left(\sum_{i=1}^{n} \lambda_i p_i\right)^* = \sum_{i=1}^{n} \overline{\lambda}_i p_i,$$

Así, vemos que $h(f) = f^*$. □

7.9.9. Ejemplo. En el Teorema 7.9.4 la hipótesis de que el espacio vectorial sea complejo es importante. En efecto, la función lineal $f : \mathbb{R}^2 \to \mathbb{R}^2$ construida en el Ejemplo 7.9.2 es normal cualquiera sea $\alpha \in \mathbb{R}$ pero no tiene ningún autovector si α no es un múltiplo entero de π. ◇

Referencias

[And98] George E. Andrews, *The theory of partitions*, Cambridge Mathematical Library, Cambridge University Press, Cambridge, 1998. Reprint of the 1976 original. MR1634067 ↑154

[Bre06] Frédéric Brechenmacher, *Histoire du théorème de Jordan de la décomposition matricielle (1870–1930)*, Tesis de doctorado, École des Hautes Études en Sciences sociales, París, Francia, 2006. ↑158

[Bru87] Richard A. Brualdi, *The Jordan canonical form: an old proof*, Amer. Math. Monthly **94** (1987), no. 3, 257–267, DOI 10.2307/2323392. MR883292 ↑158

[Cat62] S. Cater, *An Elementary Development of the Jordan Canonical Form*, Amer. Math. Monthly **69** (1962), no. 5, 391–393, DOI 10.2307/2312130. MR1531677 ↑158

[Cay58] Arthur Cayley, *A Memoir on the Theory of Matrices*, Phil. Trans. R. Soc. of London **148** (1858), 17-37, disponible en `http://www.jstor.org/stable/108649`. ↑129

[FS83] Roger Fletcher y Danny C. Sorensen, *An algorithmic derivation of the Jordan canonical form*, Amer. Math. Monthly **90** (1983), no. 1, 12–16, DOI 10.2307/2975686. MR691009 ↑158

[Fil71] Alekseĭ Fedorovich Filippov, *A short proof of the theorem on reduction of a matrix to Jordan form.*, Vestnik Moskov. Univ. Ser. I Mat. Meh. **26** (1971), no. 2, 18–19 (Russian, with English summary). MR0279113 ↑158

[GW80] Anatoly M. Galperin y Zeev Waksman, *An elementary approach to Jordan theory*, Amer. Math. Monthly **87** (1980), no. 9, 728–732. MR602830 ↑158

[GG96] Israel Gohberg y Seymour Goldberg, *A simple proof of the Jordan decomposition theorem for matrices*, Amer. Math. Monthly **103** (1996), no. 2, 157–159, DOI 10.2307/2975110. MR1375060 ↑158

[Ham53] William Rowan Hamilton, *Lectures on Quaternions*, Hodges and Smith, Dublin, Ireland, 1853. Containing a Systematic Statement of a New Mathematical Method; of which the Principles Were Communicated in 1843 to the Royal Irish Academy; and which Has Since Formed the Subject of Successive Courses of Lectures, Delivered in 1848 and Subsequent Years, in the Halls of Trinity College, Dublin: with Numerous Illustrative Diagrams, and with Some Geometrical and Physical Applications. ↑129

[Jor70] Camille Jordan, *Traité des substitutions et des équations algébriques*, Gauthier-Villars, Paris, Francia, 1870. ↑156, 158

[JvN35] Pascual Jordan y John von Neumann, *On inner products in linear, metric spaces*, Ann. of Math. (2) **36** (1935), no. 3, 719–723. MR1503247 ↑167

[Noe26] Emmy Noether, *Abstrakter Aufbau der Idealtheorie in algebraischen Zahl und Funktionenkörpern*, Math. Ann. **96** (1926), 26–61. ↑51

[Roi99] Moshe Roitman, *A short proof of the Jordan decomposition theorem*, Linear and Multilinear Algebra **46** (1999), no. 3, 245–247, DOI 10.1080/03081089908818616. MR1708591 ↑158

[Väl86] Hannu Väliaho, *An elementary approach to the Jordan form of a matrix*, Amer. Math. Monthly **93** (1986), no. 9, 711–714, DOI 10.2307/2322285. MR863972 ↑158

[vdW30] Bartel Leendert van der Waerden, *Moderne Algebra. Bd. I. Unter Benutzung von Vorlesungen von E. Artin und E. Noether*, Die Grundlehren der mathematischen Wissenschaften in Einzeldarstellungen mit besonderer Berücksichtigung der Anwendungsgebiete, vol. 23, J. Springer, Berlin, 1930. ↑51

[Yca13] Bernard Ycart, *A case of mathematical eponymy: the Vandermonde determinant*, Rev. Histoire Math. **19** (2013), no. 1, 43–77, disponible en `http://arxiv.org/abs/1204.4716`. MR3155603 ↑99